



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 113979

TO: Minh-Tam Davis
Location: Rem 3a24
Friday, February 13, 2004
Art Unit: 1642
Phone: 272-0830
Serial Number: 10 / 023182

From: Jan Delaval
Location: Biotech-Chem Library
Rem 1A51
Phone: 272-2504

jan.delaval@uspto.gov

Search Notes

pretty close:
Seq #1 4-6
9/15/97

Seq #1
10/96

STIC-Bjotech/ChemLib

113979

From: Chan, Christina
Sent: Tuesday, February 10, 2004 12:58 PM
To: Davis, Minh-Tam; STIC-Biotech/ChemLib
Subject: RE: Rush search request for 10/023182

CRIE

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

-----Original Message-----

From: Davis, Minh-Tam
Sent: Tuesday, February 10, 2004 12:56 PM
To: Chan, Christina
Subject: Rush search request for 10/023182

Please search in commercial database, issued patent files, PGPUB and interference:

- 1) SEQ ID NO:1.
- 2) The amino acid sequence encoded by SEQ ID NO:1, as shown in the sequence listing.
- 3) SEQ ID NO:4, 5,6 with and without size limitation to the size of the corresponding sequence.

Thank you.

MINH TAM DAVIS
ART UNIT 1642,
RESEM ROOM 3A24, MB 3C18
272-0830

Searcher: [Signature]
Phone: 22504
Location: _____
Date Picked Up: 2/10
Date Completed: 2/13
Searcher Prep/Review: _____
Clerical: 15
Online time: md

TYPE OF SEARCH:
NA Sequences: ✓
AA Sequences: ✓
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: ✓
WWW/Internet: _____
Other (specify): _____

ATTORNEY/AGENT INFORMATION:
 NAME: Hanson, No. 6252052man D.
 REGISTRATION NUMBER: 30,946
 REFERENCE/DOCKET NUMBER: LUD 5466.3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 688-9200
 TELEFAX: (212) 838-3884
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 11 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-09-062-422-4

Query Match 100.0%; Score 51; DB 3; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.0081;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLMMITQC 9
 |||||
 Db 1 SLMMITQC 9

Search completed: February 10, 2004, 17:20:31
 Job time : 11.7931 secs

```

; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Homo Sapiens source
US-09-574-749B-34

Query Match          100.0%; Score 51; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY      1 SLMMWITOC 9
        |||||
        1 SLMMWITOC 9

RESULT 13
US-09-183-931-42
; Sequence 42, Application US/09183931C
; Patent No. 6210886
; GENERAL INFORMATION:
; APPLICANT: Van Baren, Nicolas
; APPLICANT: Brasseur, Francis
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: METHOD FOR DIAGNOSING LEUKEMIA BY DETERMINING
; TITLE OF INVENTION: TUMOR REJECTION ANTIGEN PRECURSORS
; FILE REFERENCE: LUD 5527.1-JEL/88
; CURRENT APPLICATION NUMBER: US/09/183,931C
; EARLIER FILING DATE: 2000-02-28
; EARLIER APPLICATION NUMBER: US 09/018,422
; NUMBER OF SEQ ID NOS: 44
; SEQ ID NO 42
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-09-183-931-42

Query Match          100.0%; Score 51; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0081;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY      1 SLMMWITOC 9
        |||||
        1 SLMMWITOC 9

RESULT 14
US-09-359-503-4
; Sequence 4, Application US/09359503
; Patent No. 6251603
; GENERAL INFORMATION:
; APPLICANT: Jager, Elke; Stockert, Elisabeth; Old, Lloyd J;
; APPLICANT: Knuth, Alexander
; TITLE OF INVENTION: Method for Determining Status of A
; TITLE OF INVENTION: Cancerous Condition By Determining Antibodies
; TITLE OF INVENTION: a Patient Sample
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Fulbright & Jaworski, L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; FILING DATE: 03-October-1996
```

```

; APPLICATION NUMBER: US/09/359,503
; FILING DATE: July 23, 1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/165,546
; FILING DATE: October 2, 1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/062,422
; FILING DATE: April 17, 1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/937,263
; FILING DATE: September 15, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/752,182
; FILING DATE: 03-October-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6251603man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5466.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-359-503-4

Query Match          100.0%; Score 51; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0081;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY      1 SLMMWITOC 9
        |||||
        1 SLMMWITOC 9

RESULT 15
US-09-062-422-4
; Sequence 4, Application US/09062422
; Patent No. 6252052
; GENERAL INFORMATION:
; APPLICANT: Stockert, Elisabeth; Jager, Elke;
; APPLICANT: Chen, Yao-Tseng; Scanlan, Matthew; Knuth, Alexander;
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: ANTIBODIES WHICH BIND TO NY-ESO-1
; TITLE OF INVENTION: CANCER ASSOCIATED PROTEINS, USES THEREOF, TRUNCATED
; TITLE OF INVENTION: FORMS OF NY-ESO-1, AND HLA BINDING PEPTIDES DERIVED
; THEREFROM
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/062,422
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/937,263
; FILING DATE: September 15, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/752,182
; FILING DATE: 03-October-1996
```



```
RESULT 9
US-09-670-456A-1
; Sequence 1, Application US/09670456A
; Patent No. 6506875
; GENERAL INFORMATION:
;   LUD-5668-PCT
;   APPLICANT: Gnjatic, Sacha
;   TITLE OF INVENTION: ISOLATED PEPTIDES WHICH BIND TO HLA-C MOLECULES AND USES
;   FILE REFERENCE: LUD 5668
;   CURRENT FILING DATE: 2001-09-26
;   NUMBER OF SEQ ID NOS: 5
;   SEQ ID NO 1
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-670-456A-1

Query Match
Best Local Similarity 100.0%; Score 51; DB 4; Length 9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 SLWMITQC 9
Db 1 SLWMITQC 9

RESULT 10
US-09-751-798-5
; Sequence 5, Application US/09751798
; Patent No. 6525177
; GENERAL INFORMATION:
;   APPLICANT: Stockert, Elisabeth; Jager, Elke;
;   APPLICANT: Chen, Yao-tseng; Scanlan, Matthew;
;   APPLICANT: Knuth, Alexander; Old, Lloyd J.
;   TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer
;   TITLE OF INVENTION: Associated Proteins, Uses Thereof,
;   TITLE OF INVENTION: Truncated Forms of NY-ESO-1, and HLA
;   NUMBER OF SEQUENCES: 8
;   CORRESPONDENCE ADDRESSES:
;   ADDRESS: Fulbright & Jaworski, L.L.P.
;   STREET: 666 Fifth Avenue
;   CITY: New York City
;   STATE: New York
;   COUNTRY: USA
;   ZIP: 10103
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
;   COMPUTER: IBM
;   OPERATING SYSTEM: PC-DOS
;   SOFTWARE: WordPerfect
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/751.798
;   FILING DATE:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 09/062,422
;   FILING DATE:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/725,182
;   FILING DATE: October 3, 1996
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Hanson, No. 652517man D.
;   REGISTRATION NUMBER: 30,946
;   REFERENCE/DOCKET NUMBER: LUD 5466.3
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (212) 318-3168
;   TELEFAX: (212) 752-5958
;   INFORMATION FOR SEQ ID NO: 5:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 9 amino acids
```

```
TYPE: amino acid
TOPOLOGY: linear
US-09-751-798-5

Query Match
Best Local Similarity 100.0%; Score 51; DB 4; Length 9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 SLWMITQC 9
Db 1 SLWMITQC 9

RESULT 11
US-09-344-040C-131
; Sequence 131, Application US/09344040C
; Patent No. 6548064
; GENERAL INFORMATION:
;   APPLICANT: Tureci, Ozlem
;   APPLICANT: Sahin, Ugur
;   APPLICANT: Freundschuh, Michael
;   APPLICANT: Ramensee, Hans Georg
;   APPLICANT: Stevanovic, Stefan
;   TITLE OF INVENTION: Method for Determining Presence of Cancer in a Sample By Determining Expression of an Ssx Gene, Peptides Derived From Said Ssx Gene
;   FILE REFERENCE: LUD 5556.1
;   CURRENT APPLICATION NUMBER: US/09/344,040C
;   PRIOR FILING DATE: 1999-06-25
;   PRIOR APPLICATION NUMBER: US 09/105,839
;   PRIOR FILING DATE: 1998-06-26
;   PRIOR APPLICATION NUMBER: US 08/851,130
;   PRIOR FILING DATE: 1997-05-05
;   NUMBER OF SEQ ID NOS: 132
;   SEQ ID NO 131
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-344-040C-131

Query Match
Best Local Similarity 100.0%; Score 51; DB 4; Length 9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 SLWMITQC 9
Db 1 SLWMITQC 9

RESULT 12
US-09-574-749B-34
; Sequence 34, Application US/09574749B
; Patent No. 6548299
; GENERAL INFORMATION:
;   APPLICANT: ROSENZWEIG, Michael
;   APPLICANT: PYKEIT, Mark U.
;   APPLICANT: SCADDEN, David T.
;   APPLICANT: POZNANSKY, Mark C.
;   TITLE OF INVENTION: LYMPHOID TISSUE-SPECIFIC CELL PRODUCTION
;   TITLE OF INVENTION: FROM HEMATOPOIETIC PROGENITOR CELLS IN THREE-DIMENSIONAL
;   FILE REFERENCE: C1005/7012/KA/ERG
;   CURRENT APPLICATION NUMBER: US/09/574,749B
;   PRIOR FILING DATE: 2002-05-31
;   PRIOR APPLICATION NUMBER: US 60/107,972
;   PRIOR FILING DATE: 1998-11-12
;   PRIOR APPLICATION NUMBER: PCT/US99/26795
;   PRIOR FILING DATE: 1999-11-12
;   PRIOR APPLICATION NUMBER: US 09/524,749
;   PRIOR FILING DATE: 2000-05-18
;   NUMBER OF SEQ ID NOS: 58
;   SOFTWARE: FastSeq for Windows Version 3.0
;   SEQ ID NO 34
```

Query Match	100.0%;	Score 51;	DB 3;	Length 9;
Best Local Similarity	100.0%;	Pred. No. 2.5e+05;		
Matches	9;	Conservative 0;	Mismatches 0;	Indels 0;
				Gaps 0;

```

RESULT 5
US-09-166-448-80
; Sequence 80, Application US/09166448
Patent No. 6291430
GENERAL INFORMATION:
APPLICANT: Chaux, Pascal
APPLICANT: Vanlomme, Valrie
APPLICANT: Stroobant, Vincent
APPLICANT: Boon-Falleur, Thierry
APPLICANT: van der Bruggen, Pierre
APPLICANT: Thielemans, Kris
APPLICANT: Corbals, Jurgen
TITLE OF INVENTION: PAGE3 PEPTIDES PRESENTED BY HLA CLASS II MOLECULES
FILE REFERENCE: L0461/7052
CURRENT APPLICATION NUMBER: US/09/166,448
CURRENT FILING DATE: 1998-10-05
NUMBER OF SEQ ID NOS: 81
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 80
LENGTH: 9
TYPE: PRT
; ORGANISM: Homo sapiens
US-09-166-448-80

```

Query Match	100.0%;	Score 51;	DB 3;	Length 9;
Best Local Similarity	100.0%;	Pred. No. 2.5e+05;		
Matches	9;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;

```

RESULT 6
US-09-705-160-43
Sequence 43 Application US/09705160
Patent No. 6387630
GENERAL INFORMATION:
APPLICANT: Van Baren, Nicolas
APPLICANT: Brasseur, Francis
APPLICANT: Boon-Fallour, Thierry
TITLE OF INVENTION: METHOD FOR DIAGNOSING LEUKEMIA BY DETERMINING
TITLE OF INVENTION: TUMOR REJECTION ANTIGEN PRECURSORS
FILE REFERENCE: IUD 5527.3-JBL/MAS
CURRENT APPLICATION NUMBER: US/09/705,160
CURRENT FILING DATE: 2001-11-02
PRIOR APPLICATION NUMBER: US 09/183,931
PRIOR FILING DATE: 1998 - 10 - 30
NUMBER OF SEQ ID NOS: 44
SEQ ID NO 43
LENGTH: 9
TYPE: prt
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION:
US-09-705-160-43

```

```
Query Match      100.0%; Score 51; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Db 1 SLIMWITOC 9

RESULT 7
US-09-440-621-2
; Sequence 2, Application US/09440621

Query Match	100.0%;	Score 51;	DB 4;	Length 9;
Best Local Similarity	100.0%;	Pred. No. 2.5e+05;		
Matches	9;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;

```

RESULT 8
US-09-697-884-80
; Sequence 80, Application US/09657884
; Patent No. 6426217
; GENERAL INFORMATION:
; APPLICANT: Chaux, Pascal
; APPLICANT: Vantomme, Val,rie
; APPLICANT: Strobabant, Vincent
; APPLICANT: Boon-Falleur, Thierry
; APPLICANT: van der Bruggen, Pierre
; APPLICANT: Thielemans, Kris
; APPLICANT: Corthals, Jurgen
; TITLE OF INVENTION: MAGE-3 PEPTIDES PRESENTED BY HLA CLASS II MOLECULES
; FILE REFERENCE: L0461/7052
; CURRENT APPLICATION NUMBER: US/09/697,884
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: 09/166,448
; PRIOR FILING DATE: 1998-10-05
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 80
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-697-884-80

```

Query Match	100.0%;	Score 51;	DB 4;	Length 9;
Best Local Similarity	100.0%;	Pred. No. 2.5e+05;		
Matches	9;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;

Qy 1 SLMMITOC 9
 |||||
Db 1 SLMMITOC 9

ZIP: 10103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/359,503
FILING DATE: July 23, 1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/165,546
FILING DATE: October 2, 1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/062,422
FILING DATE: April 17, 1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/937,263
FILING DATE: September 15, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/752,182
FILING DATE: 03-October-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hansen, No. 6251603man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5466.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3000
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-359-503-5

Query Match 100.0%; Score 51; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLMMWITQC 9
DB 1 SLMMWITQC 9

RESULT 3
US-09-062-422-5
Sequence 5, Application US/09062422
Patent No. 6252052
GENERAL INFORMATION:
APPLICANT: Stockert, Elisabeth; Jager, Elke;
APPLICANT: Chen, Yao-Tsang; Scanlan, Matthew; Knuth, Alexander;
APPLICANT: Old, Lloyd J.
TITLE OF INVENTION: ANTIBODIES WHICH BIND TO NY-ESO-1
TITLE OF INVENTION: CANCER ASSOCIATED PROTEINS, USES THEREOF, TRUNCATED
TITLE OF INVENTION: FORMS OF NY-ESO-1, AND HLA BINDING PEPTIDES DERIVED
TITLE OF INVENTION: THEREFROM
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/062,422
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/937,263
FILING DATE: September 15, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/752,182
FILING DATE: 03-October-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hansen, No. 6252052man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5466.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-062-422-5

Query Match 100.0%; Score 51; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLMMWITQC 9
DB 1 SLMMWITQC 9

RESULT 4
US-08-937-263B-5
Sequence 5, Application US/08937263B
Patent No. 6274145
GENERAL INFORMATION:
APPLICANT: Chen, Yao-Tsang; Scanlan, Matthew;
APPLICANT: Gure, Ali; Old, Lloyd J.; Jager, Elke;
APPLICANT: Alexander, Knuth; Ditzhout, Jan W.
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE
TITLE OF INVENTION: ENCODING CANCER ASSOCIATED ANTIGEN, THE ANTIGEN
TITLE OF INVENTION: ITSELF, AND USES THEREOF
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski, L.L.P.
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/937,263B
FILING DATE: September 15, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/725,182
FILING DATE: October 3, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Sinn, Eric, Patent Agent
REGISTRATION NUMBER: 40,177
REFERENCE/DOCKET NUMBER: LUD 5466.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3000
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-937-263B-5

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 17:14:45 / Search time 11.7931 Seconds
(without alignments)
32.290 Million cell updates/sec

Title: US-10-023-182-5

Perfect score: 51
Sequence: 1 SLWMITQC 9

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents AA:*

1: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCDUS.COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	9	3	US-09-183-931-43
2	51	100.0	9	3	US-09-359-503-5
3	51	100.0	9	3	US-09-062-422-5
4	51	100.0	9	3	US-08-937-263B-5
5	51	100.0	9	3	US-09-166-448-80
6	51	100.0	9	4	US-09-705-160-43
7	51	100.0	9	4	US-09-440-621-2
8	51	100.0	9	4	US-09-697-884-80
9	51	100.0	9	4	US-09-670-456A-1
10	51	100.0	9	4	US-09-751-798-5
11	51	100.0	9	4	US-09-344-040C-131
12	51	100.0	9	4	US-09-574-749B-34
13	51	100.0	11	3	US-09-183-931-42
14	51	100.0	11	3	US-09-359-503-4
15	51	100.0	11	3	US-09-062-422-4
16	51	100.0	11	3	US-08-937-263B-4
17	51	100.0	11	3	US-09-166-448-79
18	51	100.0	11	4	US-09-705-160-42
19	51	100.0	11	4	US-09-440-621-1
20	51	100.0	11	4	US-09-697-884-79
21	51	100.0	11	4	US-09-751-798-4
22	51	100.0	11	4	US-09-344-040C-130
23	51	100.0	11	4	US-09-574-749B-33
24	51	100.0	180	2	US-08-791-495-7
25	51	100.0	180	2	US-08-791-495-9
26	51	100.0	180	3	US-08-937-263B-8
27	51	100.0	180	4	US-09-751-798-8

28	47	92.2	10	3	US-09-359-503-7	Sequence 7, Appl
29	43	84.3	9	4	US-09-344-040C-123	Sequence 123, App
30	42	82.4	9	4	US-09-440-621-6	Sequence 6, Appl
31	42	82.4	9	4	US-09-440-621-7	Sequence 7, Appl
32	42	82.4	9	4	US-09-440-621-8	Sequence 8, Appl
33	42	82.4	9	4	US-09-440-621-10	Sequence 10, Appl
34	37	72.5	9	3	US-09-183-931-44	Sequence 44, Appl
35	37	72.5	9	3	US-09-359-503-6	Sequence 6, Appl
36	37	72.5	9	3	US-09-062-422-6	Sequence 6, Appl
37	37	72.5	9	3	US-08-937-263B-6	Sequence 6, Appl
38	37	72.5	9	3	US-09-166-448-81	Sequence 81, Appl
39	37	72.5	9	4	US-09-705-160-44	Sequence 44, Appl
40	37	72.5	9	4	US-09-440-621-3	Sequence 3, Appl
41	37	72.5	9	4	US-09-697-884-81	Sequence 81, Appl
42	37	72.5	9	4	US-09-751-798-6	Sequence 6, Appl
43	37	72.5	9	4	US-09-344-040C-122	Sequence 122, App
44	37	72.5	9	4	US-09-344-040C-132	Sequence 132, App
45	37	72.5	9	4	US-09-574-749B-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1
US-09-183-931-43
; Sequence 43, Application US/09183931C
; Parent No. 6210886
; GENERAL INFORMATION:
; APPLICANT: Van Baren, Nicolas
; APPLICANT: Brasseur, Francis
; TITLE OF INVENTION: METHOD FOR DIAGNOSING LEUKEMIA BY DETERMINING
; TITLE OF INVENTION: TUMOR REJECTION ANTIGEN PRECURSORS
; FILE REFERENCE: LUD 5527.1-JEL/ES
; CURRENT APPLICATION NUMBER: US/09/183, 931C
; CURRENT FILING DATE: 2000-02-28
; EARLIER APPLICATION NUMBER: US 09/018,422
; EARLIER FILING DATE: 1998 - 02 - 04
; NUMBER OF SEQ ID NOS: 44
; SEQ ID NO 43
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-09-183-931-43

Query Match 100.0%; Score 51; DB 3; Length 9;
Best local similarity 100.0%; Pred. No. 2.5e+05; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

Qy 1 SLWMITQC 9
Db 1 SLWMITQC 9

RESULT 2
US-09-359-503-5
; Sequence 5, Application US/09359503
; Patent No. 6251603
; GENERAL INFORMATION:
; APPLICANT: Jager, Elke; Stockert, Elisabeth; Old, Lloyd J;
; APPLICANT: Knuth, Alexander
; TITLE OF INVENTION: Method for Determining Status of A
; TITLE OF INVENTION: Cancerous Condition By Determining Antibodies
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Fulbright & Jaworski, L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA

```

; APPLICATION NUMBER: 09/062,422
; FILING DATE: April 17, 1998
; APPLICATION NUMBER: 08/937,263
; FILING DATE: September 15, 1997
; APPLICATION NUMBER: US 08/725,182
; FILING DATE: October 3, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, Norman D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 2166.4 CIP (09807811)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 318-3400
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-751-088-4

Query Match          100.0%; Score 51; DB 6; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLWMITQC 9
Db 1 SLWMITQC 9

RESULT 13
US-10-296-734-1418
; Sequence 1418, Application US/10296734
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU P07761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: Patent version 3.2
; SEQ ID NO 1418
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: NYSOla segment 11
US-10-296-734-1418

Query Match          100.0%; Score 51; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.055;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLWMITQC 9
Db 9 SLWMITQC 17

RESULT 14
US-10-296-734-1450
; Sequence 1450, Application US/10296734
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU P07761/00
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-05-26
```

```

; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: Patent version 3.2
; SEQ ID NO 1450
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: LAGEL segment 11
US-10-296-734-1450

Query Match          100.0%; Score 51; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.055;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLWMITQC 9
Db 9 SLWMITQC 17

RESULT 15
PCT-US02-29560A-330
; Sequence 330, Application PC/TUS0229560A
; GENERAL INFORMATION:
; APPLICANT: Afari, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Gish, Kurt C.
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Wilson, Keith E.
; APPLICANT: Zlotnik, Albert
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-002710PC
; CURRENT APPLICATION NUMBER: PCT/US02/29560A
; CURRENT FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: US 60/323,469
; PRIOR FILING DATE: 2001-09-17
; NUMBER OF SEQ ID NOS: 412
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 330
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
; PCT-US02-29560A-330

Query Match          100.0%; Score 51; DB 1; Length 135;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLWMITQC 9
Db 112 SLWMITQC 120

Search completed: February 10, 2004, 17:30:39
Job time : 14.348 secs
```

PRIOR FILING DATE: 1998-10-02
PRIOR APPLICATION NUMBER: US 09/062,422
PRIOR FILING DATE: 1998-04-17
PRIOR APPLICATION NUMBER: US 08/725,162
PRIOR FILING DATE: 1996-10-03
NUMBER OF SEQ ID NOS: 18
SEQ ID NO 14
LENGTH: 10
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
NAME/KEY: Mutagen
LOCATION: 10
OTHER INFORMATION: Synthesized. At position 10 of SEQ ID NO: 11, Phe is replaced by
US-10-706-475-14

Query Match 100.0%; Score 51; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLWMITOC 9
DB 1 SLWMITOC 9

RESULT 10
US-09-165-546D-4
Sequence 4, Application US/09165546D
GENERAL INFORMATION:
APPLICANT: Knuth, Alexander; Jager, Elke; Chen, Yao, Scanlan, Matt;
Gure, Ali, Old, Lloyd, Ritter, Gerd
TITLE OF INVENTION: ISOLATED PEPTIDES CORRESPONDING TO AMINO ACID
SEQUENCES OF NY-ESO-1, WHICH BIND TO MHC CLASS I AND MHC CI
USES THEREOF

NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: FULBRIGHT & JAWORSKI LLP
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10158
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Word

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/165,546D
FILING DATE: 02-Oct-1998
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/062,422
FILING DATE: April 17, 1998
APPLICATION NUMBER: 08/937,263
FILING DATE: September 15, 1997
APPLICATION NUMBER: US 08/725,162
FILING DATE: October 3, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, Norman D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 2166.4 CIP (09807811)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3000
TELEFAX: (212) 318-3400
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-165-546D-4

Query Match 100.0%; Score 51; DB 5; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLWMITOC 9
DB 1 SLWMITOC 9

RESULT 11
US-10-706-475-1
Sequence 1, Application US/10706475
GENERAL INFORMATION:
APPLICANT: Certundelo, Vincenzo
TITLE OF INVENTION: Derivatives of NY-ESO-1 peptides and Uses Thereof
FILE REFERENCE: LUD 5629.1
CURRENT APPLICATION NUMBER: US/10/706,475
CURRENT FILING DATE: 2003-11-11
PRIOR APPLICATION NUMBER: US/09/514,036E
PRIOR FILING DATE: 2000-02-25
PRIOR APPLICATION NUMBER: US 09/440,621
PRIOR FILING DATE: 1998-11-15
PRIOR APPLICATION NUMBER: US 09/165,546
PRIOR FILING DATE: 1998-10-02
PRIOR APPLICATION NUMBER: US 09/062,422
PRIOR FILING DATE: 1998-04-17
PRIOR APPLICATION NUMBER: US 08/725,162
PRIOR FILING DATE: 1996-10-03
NUMBER OF SEQ ID NOS: 18
SEQ ID NO 1
LENGTH: 11
TYPE: PRT
ORGANISM: Homo sapiens
US-10-706-475-1

Query Match 100.0%; Score 51; DB 6; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLWMITOC 9
DB 1 SLWMITOC 9

RESULT 12
US-10-751-088-4
Sequence 4, Application US/10751088
GENERAL INFORMATION:
APPLICANT: Knuth, Alexander; Jager, Elke; Chen, Yao, Scanlan, Matt;
Gure, Ali, Old, Lloyd, Ritter, Gerd
TITLE OF INVENTION: ISOLATED PEPTIDES CORRESPONDING TO AMINO ACID
SEQUENCES OF NY-ESO-1, WHICH BIND TO MHC CLASS I AND MHC C
USES THEREOF

NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: FULBRIGHT & JAWORSKI LLP
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10158
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/751,088
FILING DATE: 02-Jan-2004
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/165,546D
FILING DATE: 02-Oct-1998

US-10-751-088-5
Sequence 5, Application US/10751088
GENERAL INFORMATION:
APPLICANT: Knuth, Alexander; Jager, Elke; Chen, Yao, Scanlan, Matt;
Gute, Ali, Old, Lloyd, Ritzer, Gerald
TITLE OF INVENTION: ISOLATED PEPTIDES CORRESPONDING TO AMINO ACID
SEQUENCES OF NY-ESO-1, WHICH BIND TO MHC CLASS I AND MHC CLASS II
USERS THEREOF
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: FIDBRIGHT & JAWORSKI LLP
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10158
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/751,088
FILING DATE: 02-Jan-2004
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/165,546D
FILING DATE: 02-Oct-1998
APPLICATION NUMBER: 09/062,422
FILING DATE: April 17, 1998
APPLICATION NUMBER: 08/937,263
FILING DATE: September 15, 1997
APPLICATION NUMBER: US 08/725,182
FILING DATE: October 3, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, Norman D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 2166.4 CIP (09807811)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3000
TELEFAX: (212) 318-3400
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-751-088-5
Query Match 100.0%; Score 51; DB 6; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLWMITOC 9
DB 1 SLWMITOC 9
RESULT 7
US-10-706-475-11
Sequence 11, Application US/10706475
GENERAL INFORMATION:
APPLICANT: Cerundolo, Vincenzo
TITLE OF INVENTION: Derivatives of NY-ESO-1 Peptides and Uses Thereof
FILE REFERENCE: LUD 5629.1
CURRENT APPLICATION NUMBER: US/10/706,475
CURRENT FILING DATE: 2003-11-11
PRIOR APPLICATION NUMBER: US/09/514,036E
PRIOR FILING DATE: 2000-02-25
PRIOR APPLICATION NUMBER: US 09/440,621
PRIOR FILING DATE: 1999-11-15
PRIOR APPLICATION NUMBER: US 09/165,546
PRIOR FILING DATE: 1998-10-02

PRIOR APPLICATION NUMBER: US 09/062,422
PRIOR FILING DATE: 1998-04-17
PRIOR APPLICATION NUMBER: US 08/725,162
PRIOR FILING DATE: 1996-10-03
NUMBER OF SEQ ID NOS: 18
SEQ ID NO: 11
LENGTH: 10
TYPE: PRT
ORGANISM: Homo sapiens
US-10-706-475-11
Query Match 100.0%; Score 51; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLWMITOC 9
DB 1 SLWMITOC 9
RESULT 8
US-10-706-475-13
Sequence 13, Application US/10706475
GENERAL INFORMATION:
APPLICANT: Cerundolo, Vincenzo
TITLE OF INVENTION: Derivatives of NY-ESO-1 Peptides and Uses Thereof
FILE REFERENCE: LUD 5629.1
CURRENT APPLICATION NUMBER: US/10/706,475
CURRENT FILING DATE: 2003-11-11
PRIOR APPLICATION NUMBER: US/09/514,036E
PRIOR FILING DATE: 2000-02-25
PRIOR APPLICATION NUMBER: US 09/440,621
PRIOR FILING DATE: 1999-11-15
PRIOR APPLICATION NUMBER: US 09/165,546
PRIOR FILING DATE: 1998-10-02
PRIOR APPLICATION NUMBER: US 08/725,162
PRIOR FILING DATE: 1996-10-03
NUMBER OF SEQ ID NOS: 18
SEQ ID NO: 13
LENGTH: 10
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
NAME/KEY: Mutagen
LOCATION: 10
OTHER INFORMATION: Synthesized. At position 10, of SEQ ID NO: 11, Phe is replaced by
US-10-706-475-13
Query Match 100.0%; Score 51; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLWMITOC 9
DB 1 SLWMITOC 9
RESULT 9
US-10-706-475-14
Sequence 14, Application US/10706475
GENERAL INFORMATION:
APPLICANT: Cerundolo, Vincenzo
TITLE OF INVENTION: Derivatives of NY-ESO-1 Peptides and Uses Thereof
FILE REFERENCE: LUD 5629.1
CURRENT APPLICATION NUMBER: US/10/706,475
CURRENT FILING DATE: 2003-11-11
PRIOR APPLICATION NUMBER: US/09/514,036E
PRIOR FILING DATE: 2000-02-25
PRIOR APPLICATION NUMBER: US 09/440,621
PRIOR FILING DATE: 1999-11-15
PRIOR APPLICATION NUMBER: US 09/165,546
PRIOR FILING DATE: 1998-10-02

ZIP: 10158
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/165,546D
FILING DATE: 02-Oct-1998
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/062,422
FILING DATE: April 17, 1998
APPLICATION NUMBER: 08/937,263
FILING DATE: September 15, 1997
APPLICATION NUMBER: US 08/725,182
FILING DATE: October 3, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, Norman D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 2166.4 CIP (09807811)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3000
TELEFAX: (212) 318-3400
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-165-546D-5

Query Match 100.0%; Score 51; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLMMWITQC 9
DB 1 SLMMWITQC 9
RESULT 3
US-10-651-616-25
Sequence 25, Application US/10651616
GENERAL INFORMATION:
APPLICANT: CytoCure, LLC
APPLICANT: DURDA, PAUL
APPLICANT: KORNIICK, JAMES T.
TITLE OF INVENTION: METHODS FOR UP-REGULATING ANTIGEN EXPRESSION IN TUMORS
FILE REFERENCE: 027823-0305578
CURRENT APPLICATION NUMBER: US/10/651,616
CURRENT FILING DATE: 2003-08-29
PRIOR APPLICATION NUMBER: 60/407,492
PRIOR FILING DATE: 2002-08-29
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PatentIn version 3.1
SEQ ID NO 25
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Homo sapien artificial
US-10-651-616-25

Query Match 100.0%; Score 51; DB 6; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 4
US-10-706-475-2
Sequence 2, Application US/10706475
GENERAL INFORMATION:
APPLICANT: Cerundolo, Vincenzo
TITLE OF INVENTION: Derivatives of NY-ESO-1 Peptides and Uses Thereof
FILE REFERENCE: LUD 5629.1
CURRENT APPLICATION NUMBER: US/10/706,475
CURRENT FILING DATE: 2003-11-11
PRIOR APPLICATION NUMBER: US/09/514,036E
PRIOR FILING DATE: 2000-02-25
PRIOR APPLICATION NUMBER: US 09/440,621
PRIOR FILING DATE: 1999-11-15
PRIOR APPLICATION NUMBER: US 09/165,546
PRIOR FILING DATE: 1998-10-02
PRIOR APPLICATION NUMBER: US 09/062,422
PRIOR FILING DATE: 1998-04-17
PRIOR APPLICATION NUMBER: US 08/725,162
PRIOR FILING DATE: 1996-10-03
NUMBER OF SEQ ID NOS: 18
SEQ ID NO 2
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
US-10-706-475-2

Query Match 100.0%; Score 51; DB 6; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLMMWITQC 9
DB 1 SLMMWITQC 9
RESULT 5
US-10-289-566-24
Sequence 24, Application US/10289566
GENERAL INFORMATION:
APPLICANT: Leturcq, Didier J.
APPLICANT: Moriarty, Ann M.
APPLICANT: Jackson, Michael R.
APPLICANT: Peterson, Per A.
APPLICANT: Richards, Jon M.
TITLE OF INVENTION: A CELL THERAPY METHOD FOR THE TREATMENT OF TUMORS
FILE REFERENCE: ORT 1342CIP
CURRENT APPLICATION NUMBER: US/10/289,566
CURRENT FILING DATE: 2002-11-07
PRIOR APPLICATION NUMBER: US 60/270,252
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: US 10/080,013
PRIOR FILING DATE: 2002-02-19
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn version 3.2
SEQ ID NO 24
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic Construct
US-10-289-566-24

Query Match 100.0%; Score 51; DB 6; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLMMWITQC 9
DB 1 SLMMWITQC 9
RESULT 6

GenCore version 5.1.6
Copyright (c) 1993.- 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 17:18:11 ; Search time 13.3448 Seconds

(without alignments)
53.077 Million cell updates/sec

Title: US-10-023-182-5

Perfect score: 51

Sequence: 1 SLMMWITQC 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 378584 seqs, 7869973 residues

Total number of hits satisfying chosen parameters: 378584

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Pending Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	9	1 PCT-US03-27125-25	Sequence 25, Appl
2	51	100.0	9	5 US-09-165-546D-5	Sequence 5, Appl
3	51	100.0	9	6 US-10-651-616-25	Sequence 25, Appl
4	51	100.0	9	6 US-10-706-475-2	Sequence 2, Appl
5	51	100.0	9	6 US-10-289-566-24	Sequence 24, Appl
6	51	100.0	9	6 US-10-751-088-5	Sequence 5, Appl
7	51	100.0	10	6 US-10-706-475-11	Sequence 11, Appl
8	51	100.0	10	6 US-10-706-475-13	Sequence 13, Appl
9	51	100.0	10	6 US-10-706-475-14	Sequence 14, Appl
10	51	100.0	11	5 US-09-165-546D-4	Sequence 4, Appl
11	51	100.0	11	6 US-10-706-475-1	Sequence 1, Appl
12	51	100.0	11	6 US-10-751-088-4	Sequence 4, Appl
13	51	100.0	30	6 US-10-296-734-1418	Sequence 1418, Ap
14	51	100.0	30	6 US-10-296-734-1450	Sequence 1450, Ap
15	51	100.0	135	1 PCT-US02-29560A-330	Sequence 330, App
16	51	100.0	180	1 PCT-US02-29560A-329	Sequence 329, App
17	51	100.0	180	1 PCT-US03-27106-75	Sequence 75, Appl
18	51	100.0	180	1 PCT-US03-27106-74	Sequence 74, Appl
19	51	100.0	180	1 PCT-US03-38193-1270	Sequence 1270, Ap
20	51	100.0	180	1 PCT-US04-00885-95	Sequence 95, Appl
21	51	100.0	180	5 US-09-165-546D-15	Sequence 15, Appl
22	51	100.0	180	6 US-10-296-734-832	Sequence 832, App
23	51	100.0	180	6 US-10-296-734-834	Sequence 834, App
24	51	100.0	180	6 US-10-657-022-74	Sequence 74, Appl
25	51	100.0	180	6 US-10-657-022-75	Sequence 75, Appl
26	51	100.0	180	6 US-10-723-860-1270	Sequence 1270, Ap

27	51	100.0	180	6 US-10-751-088-15	Sequence 15, Appl
28	51	100.0	180	6 US-10-756-149-5024	Sequence 5024, Ap
29	51	100.0	3541	5 US-10-296-734-1454	Sequence 1454, Ap
30	47	92.2	10	5 US-09-165-546D-7	Sequence 7, Appl
31	47	92.2	10	6 US-10-751-088-7	Sequence 7, Appl
32	42	82.4	9	6 US-10-706-475-6	Sequence 6, Appl
33	42	82.4	9	6 US-10-706-475-7	Sequence 7, Appl
34	42	82.4	9	6 US-10-706-475-8	Sequence 8, Appl
35	42	82.4	9	6 US-10-706-475-10	Sequence 10, Appl
36	42	82.4	9	6 US-10-706-475-12	Sequence 12, Appl
37	42	82.4	9	6 US-10-289-566-25	Sequence 25, Appl
38	42	82.4	10	6 US-10-706-475-15	Sequence 15, Appl
39	42	82.4	10	6 US-10-706-475-16	Sequence 16, Appl
40	42	82.4	10	6 US-10-706-475-17	Sequence 17, Appl
41	42	82.4	10	6 US-10-706-475-18	Sequence 18, Appl
42	37	72.5	9	5 US-09-165-546D-6	Sequence 6, Appl
43	37	72.5	9	6 US-10-706-475-3	Sequence 3, Appl
44	37	72.5	9	6 US-10-289-566-26	Sequence 26, Appl
45	37	72.5	9	6 US-10-751-088-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
PCT-US03-27125-25
Sequence 25, Application PC/TUS0327125
GENERAL INFORMATION:
APPLICANT: CytoCure, LLC
APPLICANT: DURDA, PAUL
APPLICANT: KURNICK, JAMES T.
TITLE OF INVENTION: METHODS FOR UP-REGULATING ANTIGEN EXPRESSION IN TUMORS
FILE REFERENCE: 027823-0305578
CURRENT APPLICATION NUMBER: PCT/US03/27125
CURRENT FILING DATE: 2003-08-29
PRIOR APPLICATION NUMBER: 60/407,492
PRIOR FILING DATE: 2002-08-29
NUMBER OF SEQ ID NOS: 28
SOFTWARE: Patentin version 3.1
SEQ ID NO 25
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Homo sapien artificial
PCT-US03-27125-25

Query Match 100.0%; Score 51; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLMMWITQC 9
Db 1 SLMMWITQC 9

RESULT 2
US-09-165-546D-5
Sequence 5, Application US/09165546D
GENERAL INFORMATION:

APPLICANT: Knuth, Alexander; Jager, Elke; Chen, Scanlan, Matt;
Gure, Ali, Old, Lloyd, Ritzer, Gerd
TITLE OF INVENTION: ISOLATED PEPTIDES CORRESPONDING TO AMINO ACID
SEQUENCES OF NY-ESO-1, WHICH BIND TO MHC CLASS I AND MHC C
US8 THERIOF

NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: FULBRIGHT & JAWORSKI LLP
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA

CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/751,798
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/062,422
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/725,182
FILING DATE: October 3, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. US20020010321Alman D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5466.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3168
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-751-798-4

Query Match 100.0%; Score 51; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.081;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLWMITQC 9
Db 1 SLWMITQC 9

RESULT 14
US-09-766-889A-50
Sequence 50, Application US/09766889A
Patent No. US2002016454A1
GENERAL INFORMATION:
APPLICANT: Luiten, Rosalie
APPLICANT: Boon-Falleur, Thierry
APPLICANT: van der Bruggen, Pierre
APPLICANT: Stroobant, Vincent
APPLICANT: Demotte, Nathalie
APPLICANT: Schultz, Ewijn
TITLE OF INVENTION: MAGE ANTIGENIC PEPTIDES WHICH BIND HLA-B35 AND HLA-B44
FILE REFERENCE: 10461/7104
CURRENT APPLICATION NUMBER: US/09/766,889A
CURRENT FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: US 60/177,242
PRIOR FILING DATE: 2000-01-20
PRIOR APPLICATION NUMBER: US 60/243,212
PRIOR FILING DATE: 2000-10-25
NUMBER OF SEQ ID NOS: 59
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 50
LENGTH: 11
TYPE: PRT
ORGANISM: Homo sapiens
US-09-766-889A-50

Query Match 100.0%; Score 51; DB 10; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.081;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLWMITQC 9

Db 1 SLWMITQC 9

RESULT 15
US-10-170-832-79
Sequence 79, Application US/10170832
Publication No. US2003017092A1
GENERAL INFORMATION:
APPLICANT: Chaux, Pascal
APPLICANT: Vantomme, Valrie
APPLICANT: Stroobant, Vincent
APPLICANT: Boon-Falleur, Thierry
APPLICANT: van der Bruggen, Pierre
APPLICANT: Thielemans, Kris
APPLICANT: Cortbals, Jurgien
TITLE OF INVENTION: MAGE-3 PEPTIDES PRESENTED BY HLA CLASS II MOLECULES
FILE REFERENCE: 10461/7052
CURRENT APPLICATION NUMBER: US/10/170,832
CURRENT FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: US/09/166,448
PRIOR FILING DATE: 1998-10-05
NUMBER OF SEQ ID NOS: 81
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 79
LENGTH: 11
TYPE: PRT
ORGANISM: Homo sapiens
US-10-170-832-79

Query Match 100.0%; Score 51; DB 12; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.081;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLWMITQC 9
Db 1 SLWMITQC 9

Search completed: February 10, 2004, 17:32:13
JOB time : 25.1379 secs

FILING DATE: October 3, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. US20020164665Alman D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5466.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3168
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-023-182-5

Query Match 100.0%; Score 51; DB 14; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLWMITQC 9
DB 1 SLWMITQC 9

RESULT 10
US-10-274-017-1
Sequence 1, Application US/10274017
Publication No. US20030050451A1
GENERAL INFORMATION:
LUD-5668-PCT
APPLICANT: Gnjaletic, Sacha
APPLICANT: Old, Lloyd J.
TITLE OF INVENTION: ISOLATED PEPTIDES WHICH BIND TO HLA-C MOLECULES AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: LUD 5668
CURRENT APPLICATION NUMBER: US/10/274,017
CURRENT FILING DATE: 2002-10-18
PRIOR APPLICATION NUMBER: US/09/670,456
PRIOR FILING DATE: 2001-09-26
NUMBER OF SEQ ID NOS: 5
SEQ ID NO 1
LENGTH: 9
TYPE: PPT
ORGANISM: Homo sapiens
US-10-274-017-1

Query Match 100.0%; Score 51; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLWMITQC 9
DB 1 SLWMITQC 9

RESULT 11
US-10-080-013-24
Sequence 24, Application US/10080013
Publication No. US20030077248A1
GENERAL INFORMATION:
APPLICANT: Moriarty, Ann
APPLICANT: Leturcy, Didier
APPLICANT: Degraw, Juli
APPLICANT: Heiskala, Marja
APPLICANT: Peterson, Per
APPLICANT: Jackson, Michael
TITLE OF INVENTION: A CELL THERAPY METHOD FOR THE TREATMENT OF TUMORS
FILE REFERENCE: ORF-1557
CURRENT APPLICATION NUMBER: US/10/080,013
CURRENT FILING DATE: 2002-02-19
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn version 3.1

SEQ ID NO 24
LENGTH: 9
TYPE: PPT
ORGANISM: Homo sapiens
US-10-080-013-24

Query Match 100.0%; Score 51; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLWMITQC 9
DB 1 SLWMITQC 9

RESULT 12
US-10-161-097-34
Sequence 34, Application US/10161097
Publication No. US20030096404A1
GENERAL INFORMATION:
APPLICANT: ROSENZWEIG, Michael
APPLICANT: PYKETT, Mark J.
APPLICANT: SCADDEN, David T.
APPLICANT: POZNANSKY, Mark C.
TITLE OF INVENTION: LYMPHOID TISSUE-SPECIFIC CELL PRODUCTION
TITLE OF INVENTION: FROM HEMATOPOIETIC PROGENITOR CELLS IN THREE-DIMENSIONAL
FILE REFERENCE: C1005/7012/KA/ERG
CURRENT APPLICATION NUMBER: US/10/161,097
CURRENT FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: US/09/574,749
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: US 60/107,972
PRIOR FILING DATE: 1998-11-12
PRIOR APPLICATION NUMBER: PCT/US99/26795
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: US 09/524,749
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 58
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 34
LENGTH: 9
TYPE: PPT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Homo Sapiens source
US-10-161-097-34

Query Match 100.0%; Score 51; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLWMITQC 9
DB 1 SLWMITQC 9

RESULT 13
US-09-751-798-4
Sequence 4, Application US/09751798
Patent No. US20020010321A1
GENERAL INFORMATION:
APPLICANT: Stockert, Elisabeth; Jager, Elke;
APPLICANT: Chen, Yao-tseng; Scatman, Matthew;
APPLICANT: Knuth, Alexander; Old, Lloyd J.
TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer
TITLE OF INVENTION: Associated Proteins, Uses Thereof,
TITLE OF INVENTION: Truncated Forms of NY-ESO-1, and HLA
TITLE OF INVENTION: Binding Peptides Derived Therefrom
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSER: Fulbright & Jaworski, L.L.P.
STREET: 666 Fifth Avenue

LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
US-10-177-277-131

Query Match
Best Local Similarity 100.0%; Score 51; DB 12; Length 9;
Pred. No. 7.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLMMITQC 9
Db 1 SLMMITQC 9

RESULT 6
US-10-117-937-197
Sequence 197, Application US/10117937
Publication No. US2003022039A1
GENERAL INFORMATION:
APPLICANT: CTL IMMUNO THERAPIES CORP.
APPLICANT: STWARD, John, J.L.
APPLICANT: DIAMOND, David, C.
APPLICANT: LIU, Liping
APPLICANT: XIE, Zhidong
TITLE OF INVENTION: EPITOPE SEQUENCES
FILE REFERENCE: CTIMM.027A
CURRENT APPLICATION NUMBER: US/10/117,937
CURRENT FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: US 60/282,211
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: US 60/337,017
PRIOR FILING DATE: 2001-11-07
PRIOR APPLICATION NUMBER: US 60/363,210
PRIOR FILING DATE: 2002-03-07
NUMBER OF SEQ ID NOS: 602
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 197
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
US-10-117-937-197

Query Match
Best Local Similarity 100.0%; Score 51; DB 12; Length 9;
Pred. No. 7.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLMMITQC 9
Db 1 SLMMITQC 9

RESULT 7
US-10-164-078A-14
Sequence 14, Application US/10164078A
Publication No. US20030228325A1
GENERAL INFORMATION:
APPLICANT: Bilsborough, Janine
APPLICANT: Schultze, Erwin
APPLICANT: Panichelli, Christophe
APPLICANT: Van der Bruggen, Pierre
APPLICANT: Boon, Thierry
TITLE OF INVENTION: Isolated Peptides Which Bind to HLA-B*8 Molecules And Uses Thereof
FILE REFERENCE: LUD-5756
CURRENT APPLICATION NUMBER: US/10/164,078A
CURRENT FILING DATE: 2002-06-05
NUMBER OF SEQ ID NOS: 35
SEQ ID NO 14
LENGTH: 9
TYPE: PRT
ORGANISM: H. sapiens
FEATURE:
US-10-164-078A-14

Query Match
Best Local Similarity 100.0%; Score 51; DB 12; Length 9;
Pred. No. 7.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLMMITQC 9
Db 1 SLMMITQC 9

RESULT 8
US-10-164-121A-15
Sequence 15, Application US/10164121A
Publication No. US20030228308A1
GENERAL INFORMATION:
APPLICANT: Zhang, Yi
APPLICANT: Boon, Thierry
APPLICANT: Van der Bruggen, Pierre
APPLICANT: Traversari, Catia
TITLE OF INVENTION: Isolated Peptides Which Bind to HLA-Cw6 Molecules And Uses Thereof
FILE REFERENCE: LUD-5771
CURRENT APPLICATION NUMBER: US/10/164,121A
CURRENT FILING DATE: 2002-08-26
NUMBER OF SEQ ID NOS: 36
SEQ ID NO 15
LENGTH: 9
TYPE: PRT
ORGANISM: H. sapiens
FEATURE:
US-10-164-121A-15

Query Match
Best Local Similarity 100.0%; Score 51; DB 12; Length 9;
Pred. No. 7.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLMMITQC 9
Db 1 SLMMITQC 9

RESULT 9
US-10-023-182-5
Sequence 5, Application US/10023182
Publication No. US2002016465A1
GENERAL INFORMATION:
APPLICANT: Stoeckert, Elisabeth; Jager, Elke;
Chen, Yao-tseng; Scallan, Matthew;
Knuth, Alexander; Old, Lloyd J.

TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer
Associated Proteins, Uses Thereof,
Truncated Forms of NY-ESO-1, and HLA
Binding Peptides Derived Therefrom

NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski, L.L.P.
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10103

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage

COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: WordPerfect

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/023,182
FILING DATE: 17-Dec-2001

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/751,798
FILING DATE: December 29, 2000

APPLICATION NUMBER: 09/062,422
FILING DATE: <Unknown>

APPLICATION NUMBER: 08/725,182

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 318-3168
 TELEFAX: (212) 752-5958
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-09-751-798-5

Query Match 100.0%; Score 51; DB 9; Length 9;
 Best Local Similarity 100.0%; Pred. No. 7.1e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLIMWITOC 9
 |||||
 DB 1 SLIMWITOC 9

RESULT 2
 US-09-766-889A-51
 ; Sequence 51, Application US/09766889A
 ; Patent No. US2002016454A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Luiten, Rosalie
 ; APPLICANT: Boon-Falleur, Thierry
 ; APPLICANT: van der Bruggen, Pierre
 ; APPLICANT: Stroobant, Vincent
 ; APPLICANT: Demotte, Nathalie
 ; APPLICANT: Schiltz, Erwin
 ; TITLE OF INVENTION: MAGE ANTIGENIC PEPTIDES WHICH BIND HLA-B35 AND HLA-B44
 ; FILE REFERENCE: 10461/7104
 ; CURRENT FILING DATE: 2001-01-19
 ; PRIOR APPLICATION NUMBER: US 60/177,242
 ; PRIOR FILING DATE: 2000-01-20
 ; PRIOR APPLICATION NUMBER: US 60/243,212
 ; PRIOR FILING DATE: 2000-10-25
 ; NUMBER OF SEQ ID NOS: 59
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 51
 ; LENGTH: 9
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-766-889A-51

Query Match 100.0%; Score 51; DB 10; Length 9;
 Best Local Similarity 100.0%; Pred. No. 7.1e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLIMWITOC 9
 |||||
 DB 1 SLIMWITOC 9

RESULT 3
 US-10-170-832-80
 ; Sequence 80, Application US/10170832
 ; Publication No. US20030170792A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Chau, Pascal
 ; APPLICANT: Vantomme, Valérie
 ; APPLICANT: Stroobant, Vincent
 ; APPLICANT: Boon-Falleur, Thierry
 ; APPLICANT: van der Bruggen, Pierre
 ; APPLICANT: Thielemans, Kris
 ; APPLICANT: Corhals, Jurgen
 ; TITLE OF INVENTION: MAGE-3 PEPTIDES PRESENTED BY HLA CLASS II MOLECULES
 ; FILE REFERENCE: 10461/7052
 ; CURRENT FILING DATE: 2002-06-12
 ; PRIOR APPLICATION NUMBER: US/10/170,832
 ; PRIOR FILING DATE: 2002-06-12
 ; PRIOR APPLICATION NUMBER: US/09/166,448
 ; PRIOR FILING DATE: 1998-10-05

NUMBER OF SEQ ID NOS: 81
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 80
 ; LENGTH: 9
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-170-832-80

Query Match 100.0%; Score 51; DB 12; Length 9;
 Best Local Similarity 100.0%; Pred. No. 7.1e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLIMWITOC 9
 |||||
 DB 1 SLIMWITOC 9

RESULT 4
 US-10-364-614-10
 ; Sequence 10, Application US/10364614
 ; Publication No. US20030175250A1
 ; GENERAL INFORMATION:
 ; APPLICANT: JAGER, Elke
 ; APPLICANT: KNUTH, Alexander
 ; APPLICANT: OLD, Lloyd
 ; APPLICANT: Gnifatic, Sacha
 ; TITLE OF INVENTION: ISOLATED PEPTIDES WHICH BIND TO HLA MOLECULES AND USES THEREOF
 ; FILE REFERENCE: LUD 5726.1 CIP
 ; CURRENT FILING DATE: 2003-02-24
 ; PRIOR APPLICATION NUMBER: US/10/364,614
 ; PRIOR FILING DATE: 2002-02-13
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 10
 ; LENGTH: 9
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-364-614-10

Query Match 100.0%; Score 51; DB 12; Length 9;
 Best Local Similarity 100.0%; Pred. No. 7.1e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLIMWITOC 9
 |||||
 DB 1 SLIMWITOC 9

RESULT 5
 US-10-177-277-131
 ; Sequence 131, Application US/10177277
 ; Publication No. US20030185844A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tureci, Ozlem
 ; APPLICANT: Sahin, Ugur
 ; APPLICANT: Freundschuh, Michael
 ; APPLICANT: Ramensee, Hans Georg
 ; APPLICANT: Stevanovic, Stefan
 ; TITLE OF INVENTION: Method for Determining Presence of Cancer In a Sample By Determining Expression of an SSX Gene, Peptides Derived From Said SSX Gene
 ; TITLE OF INVENTION: Gene, and Uses Thereof
 ; FILE REFERENCE: LUD 5556.1
 ; CURRENT FILING DATE: 2002-06-21
 ; PRIOR APPLICATION NUMBER: US/09/177,277
 ; PRIOR FILING DATE: 2002-06-21
 ; PRIOR APPLICATION NUMBER: US/09/344,040
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: US 09/105,839
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: US 08/851,130
 ; PRIOR FILING DATE: 1997-05-05
 ; NUMBER OF SEQ ID NOS: 132
 ; SEQ ID NO 131

Fri Feb 13 14:37:39 2004

us-10-023-182-5.rapb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 17:19:51 / Search time 25.1379 Seconds
(without alignments)
74.964 Million cell updates/sec

Title: US-10-023-182-5
Perfect score: 51
Sequence: 1 SLIMWITQC 9

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 801455 seqs, 209382283 residues
Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications AA:
1: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/1/pubppa/PCRT_NEW_PUB.pep.*
3: /cgn2_6/prodata/1/pubppa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/1/pubppa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubppa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/1/pubppa/PCRTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/1/pubppa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep.*
10: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/1/pubppa/US09_NEW_PUB.pep.*
12: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep.*
13: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep.*
14: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep.*
15: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep.*
16: /cgn2_6/prodata/1/pubppa/US10_NEW_PUB.pep.*
17: /cgn2_6/prodata/1/pubppa/US60_NEW_PUB.pep.*
18: /cgn2_6/prodata/1/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	9	US-09-751-798-5	Sequence 5, Appl
2	51	100.0	9	US-09-766-889A-51	Sequence 51, Appl
3	51	100.0	9	US-10-170-832-80	Sequence 80, Appl
4	51	100.0	9	US-10-364-614-10	Sequence 10, Appl
5	51	100.0	9	US-10-177-277-131	Sequence 131, Appl
6	51	100.0	9	US-10-117-937-197	Sequence 197, Appl
7	51	100.0	9	US-10-164-078A-14	Sequence 14, Appl
8	51	100.0	9	US-10-164-121A-15	Sequence 15, Appl
9	51	100.0	9	US-10-023-182-5	Sequence 5, Appl
10	51	100.0	9	US-10-274-017-1	Sequence 1, Appl
11	51	100.0	9	US-10-080-013-24	Sequence 24, Appl
12	51	100.0	9	US-10-161-097-34	Sequence 34, Appl
13	51	100.0	11	US-09-751-798-4	Sequence 4, Appl
14	51	100.0	11	US-09-766-889A-50	Sequence 50, Appl
15	51	100.0	11	US-10-170-832-79	Sequence 79, Appl

16	51	100.0	11	US-10-364-614-9	Sequence 9, Appl
17	51	100.0	11	US-10-177-277-130	Sequence 130, Appl
18	51	100.0	11	US-10-164-078A-28	Sequence 28, Appl
19	51	100.0	11	US-10-164-121A-29	Sequence 29, Appl
20	51	100.0	11	US-10-023-182-4	Sequence 4, Appl
21	51	100.0	11	US-10-161-097-33	Sequence 33, Appl
22	51	100.0	11	US-10-117-937-35	Sequence 35, Appl
23	51	100.0	135	US-10-295-027-388	Sequence 388, Appl
24	51	100.0	180	US-09-751-798-8	Sequence 8, Appl
25	51	100.0	180	US-09-849-602-30	Sequence 30, Appl
26	51	100.0	180	US-10-026-065-3	Sequence 3, Appl
27	51	100.0	180	US-10-117-937-74	Sequence 74, Appl
28	51	100.0	180	US-10-117-937-75	Sequence 75, Appl
29	51	100.0	180	US-10-295-027-386	Sequence 386, Appl
30	51	100.0	180	US-10-023-182-8	Sequence 8, Appl
31	51	100.0	180	US-10-146-473-69	Sequence 69, Appl
32	51	100.0	180	US-10-207-655-71	Sequence 71, Appl
33	51	100.0	397	US-09-821-883-27	Sequence 27, Appl
34	44	86.3	180	US-10-364-614-14	Sequence 14, Appl
35	43	84.3	9	US-09-833-039-123	Sequence 123, Appl
36	43	84.3	9	US-10-177-277-123	Sequence 123, Appl
37	42	82.4	9	US-10-080-013-25	Sequence 25, Appl
38	38	74.5	151	US-10-264-049-2659	Sequence 2659, Appl
39	38	74.5	593	US-10-094-749-2047	Sequence 2047, Appl
40	37	72.5	9	US-09-751-798-6	Sequence 6, Appl
41	37	72.5	9	US-09-766-889A-52	Sequence 52, Appl
42	37	72.5	9	US-10-170-832-81	Sequence 81, Appl
43	37	72.5	9	US-09-833-039-122	Sequence 122, Appl
44	37	72.5	9	US-10-239-113A-466	Sequence 466, Appl
45	37	72.5	9	US-10-364-614-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-09-751-798-5
Sequence 5, Application US/09751798
Parent No. US20020010321A1
GENERAL INFORMATION:
APPLICANT: Stockert, Elisabeth; Jager, Elise;
APPLICANT: Chen, Yao-tseng; Scallan, Matthew;
APPLICANT: Knuth, Alexander; Old, Lloyd J.
TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer
TITLE OF INVENTION: Associated Proteins, Uses Thereof,
TITLE OF INVENTION: Truncated Forms of NY-ESO-1, and HLA
TITLE OF INVENTION: Binding Peptides Derived Therefrom
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSER: Fulbright & Jaworski, L.L.P.
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/751,798
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/062,422
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/725,182
FILING DATE: October 3, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. US20020010321A1man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5466.3

Fri Feb 13 14:37:41 2004

us-10-023-182-5.rspc

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta; Kidney, and Brain;
 RA Sadasivan E.T., Meng Y., Rothenberg S.P.;
 RT "Cloning and Characterization of the complementary DNA and Gene
 Encoding a Folate Binding Protein.";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF219904; AAF66225.1; -
 DR InterPro: IPR004269; Folate rec.
 DR Pfam: PF03024; Folate rec; I.
 SQ SEQUENCE 255 AA; 29457 MW; 286775FA0CD12E0F CRC64;

QY Query Match 70.6%; Score 36; DB 11; Length 255;
 Best Local Similarity 62.5%; Pred. No. 71;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 2 LMMITOC 9
 12 LMMMAEC 19

RESULT 13

QY 094G3 PRELIMINARY; PRT; 257 AA.

AC 094G3;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Hypothetical protein.
 GN OSJNAB0022116.3.
 OS Oryza sativa (Rice), and
 OS Oryza sativa (Japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530, 39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=O. sativa;
 RA Wang R.A., Friesch D., Presting G., Wood T., Yu Y., Soderlund C.,
 RA Kim H., Rambo T., Henry D., Simmons J., Wilson R., Minx P., Du H.;
 RT "Rice Genomic Sequence."
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=O. sativa (Japonica cultivar-group);
 RA Wang R.A., Yu Y., Yang T.J., Nah G., Soderlund C., Chen M., Kim H.-R.,
 RA Rambo T., Sasaki C., Henry D., Oates R., Simmons J.;
 RT "Rice Genomic Sequence."
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC091238; AAK70632.1; -
 DR EMBL; AC122146; AAM47284.1; -
 DR Gramene; O94G3; -
 KW Hypothetical protein.
 SQ SEQUENCE 257 AA; 28876 MW; E3B8054BF4739532 CRC64;

QY Query Match 70.6%; Score 36; DB 10; Length 257;
 Best Local Similarity 55.6%; Pred. No. 71;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 1 LMMITOC 9
 40 ALMMWDOC 48

RESULT 14

QY 08RU3 PRELIMINARY; PRT; 351 AA.

AC 08RU3;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Putative vacuolar ATP synthase subunit d.
 GN P0700A11.12.
 OS Oryza sativa (Japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsunoto T., Yamamoto K.;
 RT "Oryza sativa (Japonica cultivar-group) genomic DNA, chromosome 1, PAC
 clone: P0700A11.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF003300; BAB89911.1; -
 DR Gramene; O8RU3; -
 DR InterPro: IPR002843; ATPsyn AC39sub.
 DR Pfam: PF01992; VATP-synt AC39; I.
 SQ SEQUENCE 351 AA; 40784 MW; 0E36E7FBEBCED10 CRC64;

QY Query Match 70.6%; Score 36; DB 10; Length 351;
 Best Local Similarity 71.4%; Pred. No. 94;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 3 LMMITOC 9
 328 LMMISSC 334

RESULT 15

QY 09FE04 PRELIMINARY; PRT; 351 AA.

AC 09FE04;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Putative vacuolar ATP synthase subunit AC3 protein (MZN14.21).
 GN MZN14.21.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosida II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sakano H., Liu S.X., Pham P.K., Yamada K., Banh J., Btgu P., Lee J.M.,
 RA Toriumi M., Yu G., Shinzaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "Full length cDNA of gene MZN14.21 (GI:3294284)."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Shim P., Brooks S., Chao O., Chen H., Kim C., Ecker J.;
 RT "Full length cDNA sequences of Arabidopsis thaliana."
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF327534; AAG42915.1; -
 DR EMBL; AF324679; AAG40030.1; -
 DR InterPro: IPR002843; ATPsyn AC39sub.
 DR Pfam: PF01992; VATP-synt AC39; I.
 DR VARIANT 176 176 I -> M.
 SQ SEQUENCE 351 AA; 40773 MW; D574C908F86B559B CRC64;

QY Query Match 70.6%; Score 36; DB 10; Length 351;
 Best Local Similarity 71.4%; Pred. No. 94;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 3 LMMITOC 9
 328 LMMISSC 334

Search completed: February 10, 2004, 17:19:44
 Job time: 27.069 secs

SQ SEQUENCE 70 AA; 8282 MW; 1B337B4EE34A7DAA CRC64;
 Query Match 70.6%; Score 36; DB 12; Length 70;
 Best Local Similarity 50.0%; Pred. No. 23;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 LIMMITQC 9
 Db 40 LILMUSQC 47

RESULT 9
 ID 090751 PRELIMINARY; PRT; 71 AA.
 AC 090751;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Viral protein 1 (Fragment).
 GN VP1
 OS Encephalomyocarditis virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Cardiovirinae.
 OX NCBI_TaxID=12104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ITL-001/96;
 RX MEDLINE=99049632; PubMed=9833886;
 RA Knowles N.J., Dickinson N.D., Wilden G., Carra E., Brocchi E.,
 RA De Simone F.;
 RT "Molecular analysis of encephalomyocarditis viruses isolated from pigs
 RT and rodents in Italy.";
 RL Virus Res. 57:53-62(1998).
 DR EMBL, AJ006864; CA07276.1; -.
 FT NON_TER 1
 FT 71
 SQ SEQUENCE 71 AA; 8433 MW; 6GACF79961107886 CRC64;

Query Match 70.6%; Score 36; DB 12; Length 71;
 Best Local Similarity 50.0%; Pred. No. 23;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 LIMMITQC 9
 Db 41 LILMUSQC 48

RESULT 10
 ID 09AXA1 PRELIMINARY; PRT; 101 AA.
 AC 09AXA1;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE P0501G01.16 protein.
 GN P0501G01.16
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. Nipponbare;
 RA Saeki T., Matsunoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 RT clone: P0501G01.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AP002819; BAB21087.1; -.
 DR Gramene; 09AXA1; -.
 SQ SEQUENCE 101 AA; 10785 MW; E1575443B63597FE CRC64;

Query Match 70.6%; Score 36; DB 10; Length 101;

Best Local Similarity 55.6%; Pred. No. 31;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLIMMITQC 9
 Db 3 SMMWMLQC 11

RESULT 11
 ID 08UCP4 PRELIMINARY; PRT; 198 AA.
 AC 08UCP4;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DE Hypothetical protein Atu2439.
 GN Atu2439 OR AGR C 4424.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D. Sr.,
 RA Chapman P., Glendinning J., Deatherage G., Gillet M., Grant C.,
 RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
 RA Raymond C., Rouse G., Sengphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 RT C58.";
 RL Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Querillo B., Goldman B.S., Cao X., Askew M., Halling C., Millin L.,
 RA Houmel K., Gordon J., Vaudin M., Latchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Doughy D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328(2001).
 DR EMBL; AE009191; AAL43427.1; ALT_INIT.
 DR EMBL; AE008157; AAK88176.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 198 AA; 21154 MW; FA40C6548F0562F CRC64;

Query Match 70.6%; Score 36; DB 16; Length 198;
 Best Local Similarity 85.7%; Pred. No. 57;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLIMMITQC 7
 Db 25 SLIMMITQC 31

RESULT 12
 ID 09JUL02 PRELIMINARY; PRT; 255 AA.
 AC 09JUL02;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Polate binding protein.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Db 494 LQWITOC 500

RESULT 6

ID 014411 PRELIMINARY; PRT; 783 AA.

AC 014411

DT 01-JAN-1998 (TREMBLrel. 05, Created)

DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE OptiP.

GN OptiP.

OS Candida albicans (Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; mitosporic Saccharomycetales; Candida.

OX NCBI_Taxid=5476;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=106;

RA MEDLINE=97195785; PubMed=9043116;

RA Lubkowitz M.A., Hauser L., Breslav M., Naider F., Becker J.M.;

RT "An oligopeptide transport gene from *Candida albicans*.";

RL Microbiology 143:387-396(1997).

DR EMBL: U60973; AAB6628.1; -

DR InterPro: IPR004648; Oligopept_transpt.

DR InterPro: IPR004813; Tetrapept_transpt.

DR Pfam: PF03169; OPT; 1.

DR TIGRfam: TIGR00727; ISPa_OPT; 1.

DR TIGRfam: TIGR00728; OPT sfam; 1.

SO SEQUENCE 783 AA; 88406 MW; 9298C8FF81D9CD6A CRC64;

Query Match 72.5%; Score 37; DB 3; Length 783;

Best Local Similarity 75.0%; Pred. No. 1.8e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 201 LQWITOC 208

Qy 2 LQWITOC 9

Db 201 LQWITOC 208

RESULT 7

ID 09078 PRELIMINARY; PRT; 1189 AA.

AC 09078; O9X236;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE BCDNA:LD22320 protein.

GN BCDNA:LD22320 OR C97261.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephyraeidae; Drosophilidae; Drosophila.

OX NCBI_Taxid=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RA MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Cealniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,

Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

Balcer R.M., Baer A.V., Baxendale J., Bayraktarglu L., Beasley E.M.,

Besono K.Y., Bens P.V., Berman B.P., Bhandari D., Bolshakov S.,

Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,

Butler K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,

de la Paele B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

Durbin K.J., Evansgileta C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,

RA Glodet A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jatali M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Murny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Pollard J., Puri V., Reese M.G.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of *Drosophila melanogaster*.";

RL Science 287:2185-2195(2000).

RN [2]

RP SEQUENCE OF 592-1189 FROM N.A.

RA Rudin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,

RA Agbayani A., Arcaina T.T., Baxter E., Blazer R.G., Butenoff C.,

RA Champagne M., Chavez C., Chew M., Doyle C.M., Fafan D.B., Frise E.,

RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,

RA Houston K.A., Hummasti S., Kim E., Li P., Moshrefi M., Pacleb J.M.,

RA Park S., Sequeira A., Sethi H., Snit E., Svirskas R.R., Weinburg T.,

RA Cealniker S.E.;

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AE003584; AAF51300.1; -

DR EMBL: AF132194; AAD34782.1; -

DR Flybase: FBgn0027509; BCDNA:LD22320.

KM Hypothetical protein.

SO SEQUENCE 1189 AA; 134509 MW; B6869C8720A8A873 CRC64;

Query Match 72.5%; Score 37; DB 5; Length 1189;

Best Local Similarity 75.0%; Pred. No. 1.8e+02;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 270 LQWITOC 277

Qy 2 LQWITOC 9

Db 270 LQWITOC 277

RESULT 8

ID 09078 PRELIMINARY; PRT; 70 AA.

AC 09078;

DT 01-NOV-1998 (TREMBLrel. 08, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Viral protein 1 (Fragment).

GN VP1.

OS Encephalomyocarditis virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Cardioviruses.

OC NCBI_Taxid=12104;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ITL-276/90;

RA MEDLINE=99049632; PubMed=9833886;

RA Knowles N.J., Dickinson N.D., Wilsden G., Carra E., Brocchi E.,

De Simone F.;

RT "Molecular analysis of encephalomyocarditis viruses isolated from pigs

and rodents in Italy.";

RL Virus Res. 57:53-62(1998).

DR EMBL: AJ006861; CA07273.1; -

FT NON_TER 1 1

FT NON_TER 70 70

Db 157 SLMTWITQC 165

RESULT 2

Q8GUBO PRELIMINARY; PRT; 353 AA.
 AC O8GUBO;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Putative vacuolar ATPase subunit d (EC 3.6.1.34).
 GN VHA-D.
 OS Mesembryanthemum crystallinum (Common ice plant).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Caryophyllales; Caryophyllales; Aizoaceae; Mesembryanthemum.
 NC NCB1_TaxID=3544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Leaf;
 RA Kluge C., Tavares N., Goldack D., Dietz K.J.;
 RT "Cloning and expression analysis of all cDNAs coding for the vacuolar
 RT ATPase in M. crystallinum".
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ439342; CAD27914.1; --
 KW Hydrolyase.
 SQ SEQUENCE 353 AA; 40291 MW; B483F89643E056A2 CRC64;

Query Match 76.5%; Score 39; DB 10; Length 353;
 Best Local Similarity 85.7%; Pred. No. 27;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LMWITQC 9
 Db 330 LMWISQC 336

RESULT 3

O80372 PRELIMINARY; PRT; 107 AA.
 AC O80372;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE SAMPB (Fragment).
 GN SAMPB.
 OS Aster tripolium (Sea aster).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asterales; Campanulids; Asterales; Asteraceae; Asteroideae; Asteraceae;
 CC Aster.
 NC NCB1_TaxID=74787;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Brain;
 RA Uno Y., Ueno T., Shinzaki K.Y., Kanechi M., Inagaki N., Maekawa S.,
 RA Shinzaki K.;
 RT "Early Salt-Stress Effects on Expression of Genes for Aquaporin
 RT Homologues in the Halophyte Sea Aster (Aster tripolium L.)."
 RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SUPRACELLULAR LOCATION: INTERMEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE MIP/AQUAPORIN FAMILY (TC 1.A.8).
 DR EMBL; AB012268; BAA31516.1; --
 DR HSP; P47865; J34N.
 DR InterPro; IPR000425; MIP_family.
 DR Pfam; PF00230; MIP, 1.
 DR PRINTS; PR00763; MINTINSICP.
 DR Prodom; PD000295; MIP_family; 1.
 DR PROSITE; PS00221; MIP, 1.
 KW Transmembrane; Transport.
 FT NON_TER 1
 FT NON_TER 107
 SQ SEQUENCE 107 AA; 11070 MW; F39B5840CEFC7332 CRC64;

Query Match 74.5%; Score 38; DB 10; Length 107;
 Best Local Similarity 66.7%; Pred. No. 14;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLMTWITQC 9
 Db 23 SILYWIQC 31

RESULT 4

Q9UTK3 PRELIMINARY; PRT; 151 AA.
 AC Q9UTK3;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE DJ31076.1 (Novel protein) (Fragment).
 GN DJ31076.1.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Blakey S.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL035593; CAB56526.1; --
 FT NON_TER 1
 SQ SEQUENCE 151 AA; 17990 MW; 9895001E4CFAB6D6 CRC64;

Query Match 74.5%; Score 38; DB 4; Length 151;
 Best Local Similarity 85.7%; Pred. No. 19;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 LMWITQC 9
 Db 52 LMWITQC 58

RESULT 5

Q96NH3 PRELIMINARY; PRT; 593 AA.
 AC Q96NH3;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Hypothetical protein FLJ10899.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Brain;
 RA Nishi T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H.,
 RA Yamagata T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
 RA Yamamoto J., Isono Y., Kawai-Hiro Y., Saito K., Nishikawa T.,
 RA Kikuchi K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
 RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,
 RA Takahashi-Fujii A., Ohima A., Sugiyama A., Kawakami B., Suzuki Y.,
 RA Sugano S., Nagahara K., Masuno Y., Nagai K., Isogai T.;
 RT "NEDD human cDNA sequencing project."
 RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK055461; BAB70925.1; --
 KW Hypothetical protein.
 SQ SEQUENCE 593 AA; 68636 MW; B3295013BB2225BC CRC64;

Query Match 74.5%; Score 38; DB 4; Length 593;
 Best Local Similarity 85.7%; Pred. No. 65;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 LMWITQC 9

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 17:14:10 ; Search time 26.069 Seconds
(without alignments) updates/sec
89.090 Million cell updates/sec

Title: US-10-023-182-5
Perfect score: 51
Sequence: 1 SLWMITQC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	180	4 Q9Y479	Q9Y479 homo sapien
2	39	76.5	353	10 Q8CUB0	Q8CUB0 mesembryant
3	38	74.5	107	10 Q80372	Q80372 aster triplo
4	38	74.5	151	4 Q9DUX3	Q9DUX3 homo sapien
5	37	72.5	593	4 Q96NM3	Q96NM3 homo sapien
6	37	72.5	783	3 Q14411	Q14411 candida alb
7	37	72.5	1189	5 Q9VQ78	Q9VQ78 drosophila
8	36	70.6	70	12 Q90748	Q90748 encephalomy
9	36	70.6	71	12 Q90751	Q90751 encephalomy
10	36	70.6	101	10 Q9XKA1	Q9XKA1 cryza sativ
11	36	70.6	198	16 Q8UCP4	Q8UCP4 agrobacteri
12	36	70.6	255	11 Q9UL02	Q9UL02 ratelus norv
13	36	70.6	257	10 Q94G33	Q94G33 cryza sativ
14	36	70.6	351	10 Q8R033	Q8R033 cryza sativ
15	36	70.6	351	10 Q9FE04	Q9FE04 arabidopsis
16	36	70.6	379	16 Q8G4X3	Q8G4X3 bifidobacte

17	36	70.6	395	10 Q94G32	Q94G32 cryza sativ
18	36	70.6	401	11 Q8C0T0	Q8C0T0 mus musculu
19	35	68.6	155	11 Q8R001	Q8R001 mus musculu
20	35	68.6	165	11 Q8BNX7	Q8BNX7 mus musculu
21	35	68.6	209	10 Q8GWC3	Q8GWC3 arabidopsis
22	35	68.6	247	16 Q97DR8	Q97DR8 clostridium
23	35	68.6	252	10 Q9XFW5	Q9XFW5 mesembryant
24	35	68.6	261	11 Q9R0D3	Q9R0D3 cricetus
25	35	68.6	316	8 Q37535	Q37535 katharina t
26	35	68.6	423	5 Q02168	Q02168 caenorabdi
27	35	68.6	537	5 Q9W1B5	Q9W1B5 drosophila
28	35	68.6	577	10 Q9STP1	Q9STP1 arabidopsis
29	35	68.6	640	12 Q84171	Q84171 ort virus (
30	35	68.6	720	10 Q50027	Q50027 lycopersico
31	35	68.6	861	16 Q06944	Q06944 synchocyst
32	35	68.6	974	10 Q9LPV7	Q9LPV7 arabidopsis
33	35	68.6	986	5 Q9U159	Q9U159 leishmania
34	35	68.6	996	10 Q9SGP2	Q9SGP2 arabidopsis
35	35	68.6	1005	10 Q94A35	Q94A35 arabidopsis
36	34	66.7	38	16 Q8EJ47	Q8EJ47 shewanella
37	34	66.7	116	5 Q9VL54	Q9VL54 drosophila
38	34	66.7	118	16 Q8EKD6	Q8EKD6 shewanella
39	34	66.7	202	5 Q95YB6	Q95YB6 caenorabdi
40	34	66.7	234	8 Q35689	Q35689 prototheca
41	34	66.7	247	10 Q9SM46	Q9SM46 spinacia ol
42	34	66.7	251	16 Q8KE76	Q8KE76 chlorobium
43	34	66.7	299	10 Q9ZVK1	Q9ZVK1 arabidopsis
44	34	66.7	309	4 Q9NYW2	Q9NYW2 homo sapien
45	34	66.7	316	5 Q9GR21	Q9GR21 caenorabdi

ALIGNMENTS

RESULT 1
ID Q9Y479 PRELIMINARY; PRT; 180 AA.

AC Q9Y479
DT 01-NOV-1999 (TRENBER1.12, Created)
DT 01-NOV-1999 (TRENBER1.12, Last sequence update)
DT 01-OCT-2002 (TRENBER1.22, Last annotation update)
DE LAGE-15 protein (Cancer/testis antigen 2).
GN LAGE1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=99325550; PubMed=10399963;
RA Aarnoudse C.A., Van den Doel P.B., Heemkerk B., Schrier P.I.,
RT "Interleukin-2-induced, melanoma-specific T cells recognize CAMEL, an
int. J. Cancer 82:442-448(1999).
[2]
RP SEQUENCE FROM N.A.
RA Aradhyas S., Bardaro T., Galgoczy P., Yamagata T., Esposto T.,
RA Parlan H., Ciccodicola A., Kenwick S., Platzer M., D'Urso M.,
RA Nelson D.L.,
RT "Multiple pathogenic and benign genomic rearrangements occur at a 35-
kb duplication involving the NEMO and the LAGE2 genes." ;
RL Hum. Mol. Genet. 0:0-0(2001).
DR EMBL; AJ012834; CAA10194.1; -
DR EMBL; AF277315; AAL27015.1; -
SQ SEQUENCE 180 AA; 18236 MW; 9077PAP953543A25 CRC64;

Query Match 100.0%; Score 51; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.1; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SLWMITQC 9
|||||||

FT	DOMAIN	293	593	PROTEIN KINASE.
FT	NP BIND	299	307	ATP (BY SIMILARITY).
FT	BINDING	320	320	ATP (BY SIMILARITY).
FT	ACT SITE	423	423	BY SIMILARITY.
FT	CARBOHYD	49	49	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	79	79	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	133	133	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	154	154	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	669 AA;	75006 MM;	8BDD86C67C8F2A6A CRC64;

Query Match 66.7%; Score 34; DB 1; Length 669;
 Best local Similarity 66.7%; Pred. No. 95;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLWMITQC 9
 DB 496 SLVWMTLC 504

Search completed: February 10, 2004, 17:18:05
 Job time : 7.82759 secs

FT TRANSMEM 106 127 3 (POTENTIAL).
 FT DOMAIN 128 148 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 149 169 4 (POTENTIAL).
 FT DOMAIN 170 195 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 196 221 5 (POTENTIAL).
 FT DOMAIN 222 237 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 238 260 6 (POTENTIAL).
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 298 7 (POTENTIAL).
 FT DOMAIN 299 345 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 104 183 BY SIMILARITY.
 FT MOD RES 7 7 SUFFRCTION (BY SIMILARITY).
 FT MOD RES 9 9 SUFFRCTION (BY SIMILARITY).
 FT MOD RES 11 11 SUFFRCTION (BY SIMILARITY).
 FT CARBOHYD 12 12 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 345 AA; 38592 MW; BE40459F1450404A CRC64;

Query Match 66.7%; Score 34; DB 1; Length 345;
 Best Local Similarity 44.4%; Pred. No. 50;
 Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLWMWITOC 9
 Db 51 ALVWWTAC 59

RESULT 14
 MAMB_BACSM STANDARD; PRT; 513 AA.
 ID MAMB_BACSM STANDARD; PRT; 513 AA.
 AC P16639;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE Maman endo-1,4-beta-mannosidase A and B precursor (EC 3.2.1.78)
 DE (Beta-mannanase) (Endo-1,4-mannanase).
 OS Bacillus sp. (strain AM-001).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 NCBI_Taxid=1418;
 [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RX MEDLINE=90146329; PubMed=2694961;
 RA Akino T., Kato C., Horikoshi K.;
 RT "Two Bacillus beta-mannanases having different COOH termini are
 produced in Escherichia coli carrying pMAs.";
 RL Appl. Environ. Microbiol. 55:3178-3183(1989).
 CC -1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-beta-D-mannosidic
 linkages in mannan, galactomannans, glucomannans, and
 galactoglucomannans.
 CC -1- SIMILARITY: BELONGS TO FAMILY 26 OF GLYCOSYL HYDROLASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M31797; AAA22586.1; -
 DR PIR; A37219; A37219.
 DR InterPro; IPR000805; Glyco_hydro_26.
 DR Pfam; PF02156; Glyco_hydro_26; 1.
 DR PRINTS; PR00739; GHYDRASE26.
 KW Hydrolase; Glycosidase; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 513 MANNAN ENDO-1,4-BETA-MANNOSIDASE A.
 FT CHAIN 365 MANNAN ENDO-1,4-BETA-MANNOSIDASE B.
 SQ SEQUENCE 513 AA; 58430 MW; 88D105F622CDB5A8 CRC64;

Query Match 66.7%; Score 34; DB 1; Length 513;
 Best Local Similarity 66.7%; Pred. No. 73;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SLWMWITOC 9
 Db 72 SLWMWITOC 80

RESULT 15
 DAF1_CABEL STANDARD; PRT; 669 AA.
 ID DAF1_CABEL STANDARD; PRT; 669 AA.
 AC P20792;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cell-surface receptor daf-1 precursor (EC 2.7.1.37).
 GN DAF-1 OR F28C4.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderae; Caenorhabditis.
 NCBI_Taxid=6239;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90263088; PubMed=2160853;
 RA Georgi L.L., Albert P.S., Riddle D.L.;
 RT "daf-1, a C. elegans gene controlling dauer larva development,
 encodes a novel receptor protein kinase.";
 RL Cell 61:635-645(1990).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Bradshaw H., Clarke K.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 [3]
 RP REVISIONS.
 RA Mesteron R.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: DAF-1 COULD BE CELL SURFACE RECEPTOR INVOLVED IN
 CC PROCESSING ENVIRONMENTAL CUES INTO THE DEVELOPMENTAL DECISION
 CC WHETHER OR NOT TO FORM A DAUER (ENDURING) LARVA (THE DAUER
 CC LARVA IS A NONFEEDING DISPERSAL STAGE WITH GREATER RESISTANCE
 CC TO ENVIRONMENTAL STRESS).
 CC CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC TGFβ RECEPTOR SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M32877; AAA28001.1; -
 DR EMBL; AF067616; AAC19189.1; -
 DR PIR; A35103; A35103.
 DR WormPep; F28C4.1; CE17719.
 DR InterPro; IPR000472; Activin_rec.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR003605; TGFbeta_GS.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00467; GS; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW Receptor; transducer; Serine/threonine-protein kinase; ATP-binding;
 KW Transmembrane; Glycoprotein; Signal; Developmental protein.

FT SIGNAL 1 16
 FT CHAIN 17 669 CELL-SURFACE RECEPTOR DAF-1.
 FT DOMAIN 17 169 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 170 191 POTENTIAL.
 FT DOMAIN 192 669 CYTOPLASMIC (POTENTIAL).

```

DR InterPro: IPR000276; GPCR_Rhodopsin.
DR Pfam: PF05226; TAS2R_1; UNKNOWN 1.
DR ProSite, PSS0262; G_PROTEIN_RECEP_F1_2; UNKNOWN 1.
DR Receiver, G-protein coupled receptor, Transmembrane.
KM DOMAIN 1 7
FT TRANSSEM 8 28
FT DOMAIN 2 9
FT TRANSSEM 29 50
FT TRANSSEM 51 71
FT DOMAIN 72 82
FT TRANSSEM 83 103
FT DOMAIN 104 131
FT TRANSSEM 132 152
FT DOMAIN 153 184
FT TRANSSEM 185 205
FT DOMAIN 206 239
FT TRANSSEM 240 260
FT TRANSSEM 261 266
FT TRANSSEM 267 287
FT TRANSSEM 288 309
FT CARBOHYD 157 167
SQ SEQUENCE 309 AA; 35877 MW; 5EBD9AB726A52413 CRC64;
Query Match
Best Local Similarity 66.7%; Score 34; DB 1; Length 309;
Matches 6; Conservative 75.0%; Pred. No. 45;
Indels 2; Gaps 0; Mismatches 0;

DQ 2 LMMITTC 9
| | | | |
| | | | |
DQ 95 LMMITTC 102

RESULT 12
Y378_MERTH
ID Y378_MERTH STANDARD; PRT; 334 AA.
AC C26478;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 37.0 kDa protein.
GN MTH378.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
CX NCBI_TaxID=167420;
[1]
R2 SEQUENCE FROM N.A.
R3 STRAIN=Delta H;
R4 MEDLINE=98037514; PubMed=9371463;
R5 Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
R6 Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
R7 Harrison D., Hoang L., Keagle P., Lumm W., Potlter B., Qiu D.,
R8 Spadafora R., Vitare R., Wang Y., Wierzbowski J., Gibson R.,
R9 Jiwani N., Caruso A., Bush D., Sefter H., Patwell D., Prabhakar S.,
R10 McQuigall S., Shimer G., Goyal A., Petrovski S., Church G.M.,
R11 Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
R12 "Complete genome sequence of Methanobacterium thermoautotrophicum
R13 delta: functional analysis and comparative genomics.";
R14 J. Bacteriol. 179:7135-7155(1997).
CC -1 SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1 SIMILARITY: BELONGS TO THE UPF0104 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
DR EMBL: AE000823; AAB84864.1; -
DR PIR: H69148; H69148.
DR InterPro: IPR005242; Cons_hypoth374.
DR Pfam: PF03706; UPF0104; 1.

```

```

KW TIGRFAMs: TIGR00374; TIGR00374; 1
FT Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 7 27 POTENTIAL.
FT TRANSMEM 33 53 POTENTIAL.
FT TRANSMEM 120 140 POTENTIAL.
FT TRANSMEM 142 162 POTENTIAL.
FT TRANSMEM 218 238 POTENTIAL.
FT TRANSMEM 247 267 POTENTIAL.
FT TRANSMEM 277 297 POTENTIAL.
FT TRANSMEM 300 320 POTENTIAL.
SQ SEQUENCE 334 AA; 36959 MW; 46D280EC683CB611 CRC64;

Query Match 66.7%; Score 34; DB 1; Length 334;
Best Local Similarity 55.6%; Pred. No. 48;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SLIMMITQC 9
Db 223 SLISLMMEC 231

RESULT 13
CSAR_CAVPO STANDARD; PRT; 345 AA.
ID CSAR_CAVPO 070129;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DN C5a anaphylatoxin chemotactic receptor (C5a-R).
OS C5aR1.
OC Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata;
OC Mammalia; Eutheria; Rodentia; Hystriognathi; Cavidae; Cavia.
OX NCBI_TaxId=10141;
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=Hartley;
RX MEDLINE=98236152; Pubmed=9576615;
RZ Fukuda Y., Ember J.A., Yasui A., Hugli T.E.;
RT "Cloning and characterization of the guinea pig C5a anaphylatoxin
RT receptor: interspecies diversity among the C5a receptors.";
RL Int. Immunol. 10:275-283 (1998).
CC -1- FUNCTION: RECEPTOR FOR THE CHEMOTACTIC AND INFLAMMATORY PEPTIDE
CC ANAPHYLATOXIN C5A. THIS RECEPTOR STIMULATES CHEMOTAXIS, GRANULE
CC ENZYME RELEASE AND SUPEROXIDE ANION PRODUCTION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- PMW: SULFATION PLAYS A CRITICAL ROLE IN THE ASSOCIATION OF CSAR
CC WITH C5A, BUT NO SIGNIFICANT ROLE IN THE ABILITY OF THE RECEPTOR
CC TO TRANSDUCE A SIGNAL AND MOBILIZE CALCIUM IN RESPONSE TO A SMALL
CC PEPTIDE AGONIST (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/isb-sib.ch).
CC or send an email to license@isb-sib.ch.
CC -----
DR EMBL; U86103; AAC40074.1; -
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSN.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation;
KW Chemotaxis.
FT DOMAIN 1 32 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 33 55 1 (POTENTIAL).
FT DOMAIN 56 66 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 67 89 2 (POTENTIAL).
FT DOMAIN 90 105 EXTRACELLULAR (POTENTIAL).

```

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLIMIT 7
|||
Db 56 SLIMIT 62

RESULT 10

DC1_FOPPV STANDARD; PRT; 219 AA.
AC P21974;
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable deoxycytidine kinase FPOV59 (EC 2.7.1.74) (dck).
GN FPOV59 OR FPOV59 (FPOV).
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OC NCBI_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FP-1;
RA MEDLINE=90324937; PubMed=2165135;
RA Tarragha J., Winslow J., Goebel S., Johnson G.P., Taylor J.,
RA Paoletti E.;
RT "Nucleotide sequence analysis of a 10.5 kbp HindIII fragment of
RT fowlpox virus: relatedness to the central portion of the vaccinia
RT virus HindIII D region."
RT J. Gen. Virol. 71:1517-1524(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20193820; PubMed=10729156;
RA Alfonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "The genome of fowlpox virus."
RT J. Virol. 74:3815-3831(2000).
RN [3]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=94106115; PubMed=8279127;
RA Koonin E.V., Senkevich T.G.;
RT "Fowlpox virus encodes a protein related to human deoxycytidine
RT kinase: further evidence for independent acquisition of genes for
RT enzymes of nucleotide metabolism by different viruses."
RL Virus Genes 7:289-295(1993).
CC -1- CATALYTIC ACTIVITY: NTP + deoxycytidine = NDP + CMP.
CC -1- SIMILARITY: BELONGS TO THE DCK/DCK FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL, X17202; CA35067.1; -
DR EMBL, AF198100; AAF4403.1; -
DR PIR, D35216; D35216.
DR InterPro, IPR002624; DCK.
DR Pfam, PF0112; DCK.1.
KW Transferase; kinase; ATP-binding.
FT NP BIND 16
FT SEQUENCE 219 AA; 25908 MW; 4C3B08B85B55CE7 CRC64;

Query Match 66.7%; Score 34; DB 1; Length 219;
Best Local Similarity 66.7%; Fred. No. 32;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 MWITOC 9
|||
Db 179 MWITOC 184

RESULT 11

T2R8_HUMAN STANDARD; PRT; 309 AA.
AC G9NY2;
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE Taste receptor type 2 member 8 (T2R8) (taste receptor family B member
DE 5) (TRB5).
GN TAS2R8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND TOPOLOGY.
RX MEDLINE=20222571; PubMed=10761934;
RA Adler E., Hoon M.A., Mueller K.L., Chandrasekar J., Ryba N.J.P.,
RA Zuker C.S.;
RT "A novel family of mammalian taste receptors."
RT Cell 100:693-702(2000).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=20222572; PubMed=10761935;
RA Chandrasekar J., Mueller K.L., Hoon M.A., Adler E., Feng L., Guo W.,
RA Zuker C.S., Ryba N.J.;
RT "T2R8 function as bitter taste receptors."
RT Cell 100:703-711(2000).
RN [3]
RP REVIEW.
RX MEDLINE=2213574; PubMed=12139982;
RA Montmayeur J.-P., Matsunami H.;
RT "Receptors for bitter and sweet taste."
RT Curr. Opin. Neurobiol. 12:366-371(2002).
RN [4]
RP REVIEW.
RX MEDLINE=21634924; PubMed=11696554;
RA Margolske R.F.;
RT "Molecular mechanisms of bitter and sweet taste transduction."
RT J. Biol. Chem. 277:1-4(2002).
RN [5]
RP REVIEW.
RX MEDLINE=22469025; PubMed=12581520;
RA Zhang Y., Hoon M.A., Chandrasekar J., Mueller K.L., Cook B., Wu D.,
RA Zuker C.S., Ryba N.J.;
RT "Coding of sweet, bitter, and umami tastes: different receptor cells
RT sharing similar signaling pathways."
RT Cell 112:293-301(2003).
CC -1- FUNCTION: Receptor that may play a role in the perception of
CC bitterness and is gustducin-linked. May play a role in sensing the
CC chemical composition of the gastrointestinal content. The activity
CC of this receptor may stimulate alpha gustducin, mediate PLC-beta-2
CC activation and lead to the gating of TRPM5.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in subsets of taste receptor cells
CC of the tongue and palate epithelium and exclusively in gustducin-
CC positive cells.
CC -1- MISCELLANEOUS: Most taste cells may be activated by a limited
CC number of bitter compounds; individual taste cells can
CC discriminate among bitter stimuli.
CC -1- SIMILARITY: Belongs to family T2R of G-protein coupled receptors.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL, AF227134; AAF43907.1; -
DR Genbank, HGNC:14915; TAS2R8.
DR MIM, 604794; -.


```

RP SEQUENCE FROM N.A.
RC STRAIN=8288C / GRF8;
RX MEDLINE=91006011; PubMed=2209542;
RA Bartel B., Muenning I., Varshavsky A.;
RT "The recognition component of the N-end rule pathway.";
RL EMBL J. 9:3179-3189 (1990).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=8288C;
RX MEDLINE=97279231; PubMed=9133739;
RA Arroyo J., Garcia-Gonzalez M., Garcia-Saez M.I., Sanchez-Perez M.,
RA Nombela C.;
RT "DNA sequence analysis of a 23,002 bp DNA fragment of the right arm
of Saccharomyces cerevisiae chromosome VII.";
RL Yeast 13:357-363 (1997).
RN (3)
RP SEQUENCE OF 346-1950 FROM N.A.
RA Hebling U., Hofmann B., Deltus H.;
RT Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RECOGNITION COMPONENT OF THE N-END RULE PATHWAY. BINDS
CC TO PROTEINS BEARING AMINO-TERMINAL RESIDUES THAT ARE DESTABILIZING
CC ACCORDING TO THE N-END RULE, BUT DOES NOT BIND TO OTHERWISE
CC IDENTICAL PROTEINS BEARING STABILIZING AMINO-TERMINAL RESIDUES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X53747; CAA37779.1; -
DR EMBL; X59074; CAA67528.1; -
DR EMBL; Z72969; CAA97210.1; -
DR PIR; S12332; S12332.
DR SGD; S0003416; UBR1.
DR GO; GO:0005837; C:26S proteasome; IPI.
DR InterPro; IPR003126; Znf.Ntracognin.
DR InterPro; IPR001841; Znf.Ting.
DR Pfam; PF02207; zf-UBR1; 1.
DR SMART; SMO0184; RING; 1.
DR SMART; SMO0396; ZNF-UBR1; 1.
DR Ligase; Ubl conjugation pathway.
KM LIGASE; Ubl conjugation pathway.
SQ SEQUENCE 1950 AA; 224836 MW; 33B4CD3A031AF523 CRC64;

Query Match 68.6%; Score 35; DB 1; Length 1950;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Cy 2 LHMWITOC 9
Db 395 LHMWITOC 402

RESULT 8
RRPL_TUK STANDARD; PRT; 2103 AA.
ID RRPL_TUK
AC P33453;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE RNA-directed RNA polymerase (EC 2.7.7.48) (L protein).
GN L.
OS Uukuniemi virus (Tuk).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
CX NCBI_TaxID=11591;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=823;
RX MEDLINE=92333259; PubMed=1629699;
RA Elliott R.M., Dunn E., Simons J.F., Pelterson R.F.;
RT "Nucleotide sequence and coding strategy of the Uukuniemi virus L RNA

```

```

RT segment."
RL J. Gen. Virol. 73:1745-1752 (1992).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA) (N).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D10759; BAA01590.1; -
DR PIR; JQ1621; JQ1621.
DR InterPro; IPR007099; RNA_pol_NSvir.
DR Pfam; PF04196; Bunya_RdRp; 1.
KM Transferase; Nucleotidyltransferase; RNA-directed RNA polymerase.
SQ SEQUENCE 2103 AA; 241039 MW; B0BA708451B0B6BD CRC64;

Query Match 68.6%; Score 35; DB 1; Length 2103;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 2 LHMWITOC 9
Db 1207 LHMWITOC 1214

RESULT 9
WM10_YEAST STANDARD; PRT; 122 AA.
ID WM10_YEAST
AC O03880;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical 14.0 kDa protein in RPL15B-GCR3 intergenic region.
GN YMR123W OR YMR564.05.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
CX NCBI_TaxID=4932;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=8288C / AB972;
RX PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentile S., Hamlin N., Hunt S.,
RA Jagsi K., Lye G., Moule S., Odell C., Pearson D., Rajadream M.A.,
RA Rice P., Skellon J., Walsh S., Whitehead S., Barrett B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
XIII."
RL Nature 387:90-93 (1997).
CC -1- SIMILARITY: TO S.POMBE SPBP23A10.02.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z49273; CAA89272.1; -
DR PIR; S54492; S54492.
DR SGD; S0004730; PRR1.
DR Hypothetical protein; Transmembrane.
FT TRANSMEM 21
FT TRANSMEM 41
FT TRANSMEM 67
SQ SEQUENCE 122 AA; 13962 MW; 5CAAE929D55F2C40 CRC64;

Query Match 66.7%; Score 34; DB 1; Length 122;
Best Local Similarity 85.7%; Pred. No. 18;

```

```

DR 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribonucleoside-diphosphate reductase large chain (EC 1.17.4.1)
DE (Ribonucleoside reductase).
GN 028.
OS Human herpesvirus (type 7 / strain J1) (HEV7).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=57278;
RN [1]
RP SEQUENCE FROM N.A.
RA Nicholas J.
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
CC thiorodoxin + H(2)O = ribonucleoside diphosphate + reduced
CC thiorodoxin.
CC -1- PATHWAY: DNA replication pathway; first step.
CC -1- SUBUNIT: Heterodimer of a large and a small chain.
CC -1- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
CC LARGE CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U43400; AAC54690.1; -
DR PIR; T41930; T41930.
DR InterPro: IPR000788; Ribonucleo red.
DR Pfam; PF00317; ribonuc_red_1g, 1.
DR Pfam; PF02867; ribonuc_red_1gc, 1.
DR PROSITE; PS00089; RIBOFED_LARGE, 1.
DR KW Oxidoreductase; DNA replication.
SO SEQUENCE 806 AA; 93069 MW; 4BA516084FCCDC0 CRC64;

Query Match 68.6%; Score 35; DB 1; Length 806;
Best Local Similarity 66.7%; Pred. NO. 76;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLIMWITQC 9
DB 309 SLIMWITDFC 317

RESULT 6
ID CUT4_SCHPO STANDARD; PRT; 1458 AA.
AC Q9URV2; O13457;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE CUT4 protein.
GN CUT4 OR SPB106.09.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RA MEDLINE=97078755; PubMed=8918880;
RA Yamashita Y.M., Nakaseko Y., Samejima I., Kumada K., Yamada H.,
RA Michaelson D., Yanagida M.;
RT "20S cyclosoome complex formation and proteolytic activity inhibited by
RT the CAM/PKX pathway";
RT Nature 384:276-279(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=9712;

```

```

RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth N., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones S., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Oell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Gymnopoulos B.,
RA Welter J., Vanstreels B., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leinrich H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leleau V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garçon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez U., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.L.,
RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin U.,
RA Shipakovski G.V., Ussery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe";
RL Nature 415:871-880(2002).
CC -1- FUNCTION: Has an essential role in the regulation of 20S cyclosoome
CC complex formation. Mutations to this protein prevent the exit from
CC mitosis.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D85196; BAA22618.1; -
DR PIR; A110295; CAB53725.1; -
DR PIR; T39266; T39266.
DR PIR; T51895; T51895.
DR Genes; Spombe; SPB106.09; -
DR Pfam; PF01851; PC_rep; 4.
DR KW Mitosis.
FT CONFLICT 995 995 A -> V (IN REF. 1).
SO SEQUENCE 1458 AA; 165409 MW; 182E65BD3A62183 CRC64;

Query Match 68.6%; Score 35; DB 1; Length 1458;
Best Local Similarity 55.6%; Pred. NO. 1.4e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLIMWITQC 9
DB 538 SLIMWITFC 546

RESULT 7
ID UBR1_YEAST STANDARD; PRT; 1950 AA.
AC P19812;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE N-end-recognition protein (Ubiquitin-protein ligase E3 component) (N-
DE recognition).
GN UBR1 OR PTRI OR YGR184C OR G7168.
OS Schizosaccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]

```



```

FT DOMAIN 5 82 GLY-RICH.
PT TRASNEM 156 172 POTENTIAL.
SQ SEQUENCE 180 AA; 17992 MW; B122C5C2C8BE1569 CRC64;

Query Match
Best Local Similarity 100.0%; Score 51; DB 1; Length 180;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLIMMITOC 9
DB 157 SLIMMITOC 165

RESULT 2
V0D1 ARATH STANDARD; PRT; 351 AA.
ID V0D1 ARATH
AC Q9LJ15;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable vacuolar ATP synthase subunit d 1 (EC 3.6.3.14) (V-ATPase d
subunit 1) (Vacuolar proton pump d subunit 1).
GN A13G28710 OR MN14.21.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016720; PubMed=11130713;
RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unseid M.,
RA Salanoubat M., Lemcke K., Rieger M., Perez-Alonso M., Obermaier B.,
RA Fattmann B., Valle G., Bloeker H., Maché R., Puidomenech P.,
RA Deleney M., Boutry M., Grivell L.A., Maché R., Puidomenech P.,
RA De Simone V., Choise N., Artiguenave F., Robert C., Brottier F.,
RA Wincker P., Catolico L., Weissbach J., Saurin W., Quetier F.,
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Wurmbech E., Drzonek H., Erfle H., Jordan N., Bangert S.,
RA Weidelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B.,
RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,
RA Reichelt J., Scharte M., Schoen O., Bagues M., Terol J., Climent J.,
RA Navarero P., Collado C., Perez-Perez A., Ottemwelder B., Duchemin D.,
RA Cooke R., Maarse A.C., Alcaraz J.-P., Cottet A., Casasuberta E.,
RA De Haan M., Maarse A.C., Flores M., Liguori R., Vitale D.,
RA Monfort A., Argirou A., Flores M., Liguori R., Vitale D.,
RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.Y., Jenkins J.,
RA Rooney T., Rizzo W., Maite R., Wu D., Peterson J., Van Aken S.,
RA Creasy T.H., Haas B., Maite R., Wu D., Peterson J., Van Aken S.,
RA Pal G., Miltischer J., Selliers P., Gill J.E., Feldblyum T.V.,
RA Preuss D., Lin X., Nieman W.C., Salzberg S.L., White O., Venter J.C.,
RA Sasamoto S., Kimura T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsumoto A., Muraki A.,
RA Nakayama S., Nakazaki N., Shino S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.,
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
thaliana."
RL Nature 408:820-822(2000).
CC -1- FUNCTION: SUBUNIT OF THE INTEGRAL MEMBRANE V0 COMPLEX OF VACUOLAR
ATPASE. VACUOLAR ATPASE IS RESPONSIBLE FOR ACIDIFYING A VARIETY OF
INTRACELLULAR COMPARTMENTS IN EUKARYOTIC CELLS, THUS PROVIDING
MOST OF THE ENERGY REQUIRED FOR TRANSPORT PROCESSES IN THE
VACUOLAR SYSTEM.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
H(+) (Out).
CC -1- SUBUNIT: V-ATPase is a heteromultimeric enzyme composed of a
peripheral catalytic V1 complex (components A to H) attached to
an integral membrane V0 proton pore complex (components: a, c, c',
c'', and d).
CC -1- SIMILARITY: Belongs to the V-ATPase V0D/AC39 subunit family.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP000420; BAB02186.1; -
DR InterPro; IPR002843; ATPase_ATC39sub.
DR Pfam; PF01992; V-ATP-lynt_AC39; 1.
KM Hydrolyase; Hydrogen ion transport.
SQ SEQUENCE 351 AA; 40791 MW; 5E1A67A1A9AC4EF4 CRC64;

Query Match
Best Local Similarity 70.6%; Score 36; DB 1; Length 351;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 LMMITOC 9
DB 328 LMMITOC 334

RESULT 3
V0D2 ARATH STANDARD; PRT; 351 AA.
ID V0D2 ARATH
AC Q9LHA4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable vacuolar ATP synthase subunit d 2 (EC 3.6.3.14) (V-ATPase d
subunit 2) (Vacuolar proton pump d subunit 2).
GN A13G28715 OR MN14.22.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016720; PubMed=11130713;
RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unseid M.,
RA Salanoubat M., Lemcke K., Rieger M., Perez-Alonso M., Obermaier B.,
RA Fattmann B., Valle G., Bloeker H., Maché R., Puidomenech P.,
RA Deleney M., Boutry M., Grivell L.A., Maché R., Puidomenech P.,
RA De Simone V., Choise N., Artiguenave F., Robert C., Brottier F.,
RA Wincker P., Catolico L., Weissbach J., Saurin W., Quetier F.,
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Wurmbech E., Drzonek H., Erfle H., Jordan N., Bangert S.,
RA Weidelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B.,
RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,
RA Reichelt J., Scharte M., Schoen O., Bagues M., Terol J., Climent J.,
RA Navarero P., Collado C., Perez-Perez A., Ottemwelder B., Duchemin D.,
RA Cooke R., Maarse A.C., Alcaraz J.-P., Cottet A., Casasuberta E.,
RA De Haan M., Maarse A.C., Flores M., Liguori R., Vitale D.,
RA Monfort A., Argirou A., Flores M., Liguori R., Vitale D.,
RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.Y., Jenkins J.,
RA Rooney T., Rizzo W., Maite R., Wu D., Peterson J., Van Aken S.,
RA Creasy T.H., Haas B., Maite R., Wu D., Peterson J., Van Aken S.,
RA Pal G., Miltischer J., Selliers P., Gill J.E., Feldblyum T.V.,
RA Preuss D., Lin X., Nieman W.C., Salzberg S.L., White O., Venter J.C.,
RA Sasamoto S., Kimura T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsumoto A., Muraki A.,
RA Nakayama S., Nakazaki N., Shino S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.,
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
thaliana."
RL Nature 408:820-822(2000).
CC -1- FUNCTION: SUBUNIT OF THE INTEGRAL MEMBRANE V0 COMPLEX OF VACUOLAR

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 17:13:40 ; Search time 6.82759 Seconds
(without alignments)
61.990 Million cell updates/sec

Title: US-10-023-182-5
Perfect score: 51
Sequence: 1 SLMTWITQC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	180	1 CTG1_HUMAN	P78358 homo sapien
2	36	70.6	351	1 VOD1_ARATH	Q91h35 arabidopsis
3	36	70.6	351	1 VOD2_ARATH	Q91h34 arabidopsis
4	35	68.6	693	1 RAB1_SCHPO	Q96t17 schizosacch
5	35	68.6	806	1 RIR1_HSV7J	P50641 human herpes
6	35	68.6	1458	1 CUT4_SCHPO	Q9UR72 schizosacch
7	35	68.6	1950	1 UBRI_YEAST	P19812 saccharomyc
8	35	68.6	2103	1 RRPL_UOK	P13453 uukuniemi v
9	34	66.7	122	1 YML0_YEAST	Q03880 saccharomyc
10	34	66.7	219	1 DCK1_FOWPV	P21974 fowlpox vir
11	34	66.7	309	1 T2R8_HUMAN	Q9NY42 homo sapien
12	34	66.7	334	1 Y378_METHH	O26478 methanobact
13	34	66.7	345	1 CSAR_CAYPO	O70129 cavia porce
14	34	66.7	513	1 MAMB_BACSM	P16659 bacillus sp
15	34	66.7	669	1 DAF1_GAEEL	P20792 caenorhabdi
16	33	64.7	237	1 R33A_HUMAN	Q14088 homo sapien
17	33	64.7	237	1 R33A_MOUSE	P97980 mus musculu
18	33	64.7	255	1 FOL1_MOUSE	P55846 mus musculu
19	33	64.7	274	1 ATP6_BUCAI	P57118 buchnera ap
20	33	62.7	132	1 TVC3_MOUSE	P06324 mus musculu
21	32	62.7	135	1 TVC1_MOUSE	P01740 mus musculu
22	32	62.7	145	1 YDB4_MERVA	Q68749 methanococ
23	32	62.7	250	1 DIP_ANTVA	P33560 antitritinin
24	32	62.7	365	1 Y255_MYCOE	P47345 mycoplasma
25	32	62.7	473	1 KNOB_PLAFA	P13817 plasmodium
26	32	62.7	509	1 VMT9_MXYVL	P08073 myxoma viru
27	32	62.7	634	1 KNOB_PLAFA	P09346 plasmodium
28	32	62.7	657	1 KNOB_PLAFA	P09346 plasmodium
29	32	62.7	724	1 P85A_BOVIN	P23777 bos taurus
30	32	62.7	724	1 P85A_HUMAN	P23796 homo sapien
31	32	62.7	724	1 P85A_MOUSE	P26450 mus musculu
32	32	62.7	724	1 P85A_RAT	O63787 rattus norv
33	32	62.7	811	1 RFX1_YEAST	P48743 saccharomyc

ALIGNMENTS

RESULT 1	ID	CTG1_HUMAN	STANDARD;	PRT;	180 AA.
AC	P78358;				
DT	15-DEC-1998 (Rel. 37, Created)				
DT	15-DEC-1998 (Rel. 37, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Cancer/testis antigen 1 (Autoimmunogenic cancer/testis antigen NY-ESO-1).				
GN	CTAG1 OR CTAG.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_Taxid=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=97203161; PubMed=9050879;				
RA	Chen Y.-T., Scanlan M.U., Sahin U., Tuerci O., Gure A.O., Tsang S.,				
RT	Williamson B., Stockert B., Pfrendershub M., Old L.J.;				
RT	"A testicular antigen aberrantly expressed in human cancers detected				
RT	by autologous antibody screening."				
RL	Proc. Natl. Acad. Sci. U.S.A. 94:1914-1918(1997).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Melanoma;				
RX	MEDLINE=98289662; PubMed=9626360;				
RA	Lethe B., Lucas S., Michaux L., de Smet C., Godelaine D., Serrano A.,				
RT	Schwartzenruber D.J., Rosenberg S.A.;				
RT	"A breast and melanoma-shared tumor antigen: T cell responses to				
RT	antigenic peptides translated from different open reading frames."				
RL	J. Immunol. 161:3596-3606(1998).				
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN TESTIS AND OVARY AND IN A WIDE				
CC	VARIETY OF CANCERS. DETECTED IN UTERINE MYOMETRIUM.				
CC	-1- SIMILARITY: BELONGS TO THE CTAG FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL collaboration				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
DR	EMBL; U87459; AA849693.1; -				P31352 marburg vir
DR	EMBL; AJ003149; CA005908.1; -				P35262 rattus norv
DR	EMBL; AF038567; AAD05202.1; -				Q63269 rattus norv
DR	Genew; HGNC:2491; CTAG1.				Q14573 homo sapien
DR	MIM; 300156; -				P47105 saccharomyc
DR	Transmembrane; Antigen.				Q43844 solanum tub
					O47479 loligo blec
					O97665 macaca mula
					P24890 caenorhabdi
					Q35522 phytothor
					P49355 bos taurus
					P49356 homo sapien

C/Accession: T51995
 R/Yamashita, Y.; Nakaseko, Y.; Samejima, I.; Kumada, K.; Yamada, H.; Yanagida, M.
 Nature 384, 276-279, 1996
 A/Title: 20S cyclosome complex formation and proteolytic activity inhibited by the CAMP/
 A/Reference number: Z55896
 A/Accession: T51995
 A/Status: preliminary; translated from GB/EMBL/DBD
 A/Molecule type: DNA
 A/Residues: 1-1458 <YAM>
 A/Cross-references: EMBL:D85196; PDB:BA22618.1
 C/Genetics:
 A/Gene: cut4+

Query Match 68.6%; Score 35; DB 2; Length 1458;
 Best Local Similarity 55.6%; Pred. No. 2.4e+02;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SLMMITOC 9
 Db 538 SLOWITIC 546

RESULT 13

probable cell cycle control protein cut4 [similarity] - fission yeast (Schizosaccharomyces pombe)
 C/Species: Schizosaccharomyces pombe
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 20-Jun-2000
 C/Accession: T39266
 R/Lyte, M.; Rajandream, M.A.; Barrell, B.G.; Davis, P.; Churcher, C.M.
 submitted to the EMBL Data Library, August 1999
 A/Reference number: Z21840
 A/Accession: T39266
 A/Status: preliminary; translated from GB/EMBL/DBD
 A/Molecule type: DNA
 A/Residues: 1-1458 <LYN>
 A/Cross-references: EMBL:AL110295; PDB:CB53725.1; GSPDB:GN00067; SPDB:SPBC106.09
 A/Experimental source: strain 972h-; cosmid c106
 C/Genetics:
 A/Gene: SPDB:SPBC106.09
 A/Map position: 2

Query Match 68.6%; Score 35; DB 2; Length 1458;
 Best Local Similarity 55.6%; Pred. No. 2.4e+02;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SLMMITOC 9
 Db 538 SLOWITIC 546

RESULT 14

ubiquitin-protein ligase (EC 6.3.2.19) - yeast (Saccharomyces cerevisiae) (strain S288C)
 N/Alternate names: protein G7168; protein YGR184c
 C/Species: Saccharomyces cerevisiae
 C/Date: 16-Sep-1997 #sequence_revision 16-Sep-1992 #text_change 03-Jun-2002
 C/Accession: S12332; S64502; S64498
 R/Bartel, B.; Wuening, I.; Varshavsky, A.
 EMBO J. 9, 3179-3189, 1990
 A/Title: The recognition component of the N-end rule pathway.
 A/Reference number: S12332; PMID:91006011; PMID:2209542
 A/Accession: S12332
 A/Molecule type: DNA
 A/Residues: 1-1950 <BAR>
 A/Cross-references: EMBL:X53747; NID:G4743; PDB:CAA37779.1; PID:G4744
 A/Experimental source: strain S288C
 R/Arroyo, J.; Garcia-Gonzalez, M.; Garcia-Saez, M.I.; Sanchez-Perez, M.; Nomela, C.
 submitted to the Protein Sequence Database, May 1996
 A/Reference number: S64499
 A/Accession: S64502
 A/Molecule type: DNA
 A/Residues: 1-1950 <ARR>
 A/Cross-references: EMBL:Z72969; NID:G132325; PDB:CAA97210.1; PID:G132326; MIPS:YGR18

A/Experimental source: strain S288C
 R/Rehling, U.; Hofmann, B.; Deltus, H.
 submitted to the Protein Sequence Database, May 1996
 A/Reference number: S64003
 A/Accession: S64498
 A/Molecule type: DNA
 A/Residues: 1615-1950 <HEB>
 A/Cross-references: EMBL:Z72969; MIPS:YGR184c
 A/Experimental source: strain S288C
 C/Genetics:
 A/Gene: SGD:UBR1; PTR1
 A/Cross-references: SGD:S0003416; MIPS:YGR184c
 A/Map position: 7R
 C/Superfamily: ubiquitin-protein ligase
 C/Keywords: ligase; protein degradation

Query Match 68.6%; Score 35; DB 2; Length 1950;
 Best Local Similarity 50.0%; Pred. No. 3.1e+02;
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMMITOC 9
 Db 395 ILMWITIC 402

RESULT 15

genome polyprotein - Uukuniemi virus (strain S23)
 N/Alternate names: L protein S23
 N/Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
 C/Species: Uukuniemi virus
 C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jun-2000
 C/Accession: J01621
 R/Elloit, R.M.; Dunn, E.; Simons, J.F.; Pettersson, R.F.
 J. Gen. Virol. 73, 1745-1752, 1992
 A/Title: Nucleotide sequence and coding strategy of the Uukuniemi virus L RNA segment.
 A/Reference number: J01621; M01D:9233259; PMID:1629699
 A/Accession: J01621
 A/Molecule type: genomic RNA
 A/Residues: 1-2103 <ELL>
 A/Cross-references: DDBJ:D10759; NID:G222688; PDB:BA01590.1; PID:G222689
 C/Genetics:
 A/Map position: segment L
 C/Superfamily: Uukuniemi virus RNA-directed RNA polymerase
 C/Keywords: nucleotidyltransferase

Query Match 68.6%; Score 35; DB 1; Length 2103;
 Best Local Similarity 75.0%; Pred. No. 3.3e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LMMITOC 9
 Db 1207 LRMWITAC 1214

Search completed: February 10, 2004, 17:29:46
 Job time : 11.8621 secs

A:Map position: 2
A:Introns: 75/3

Query Match
Best Local Similarity 68.6%; Score 35; DB 2; Length 612;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LMMITOC 9
DB 4 IAMITOC 10

RESULT 8

T41930
Ribonucleotide reductase, large subunit - human herpesvirus 7 (strain J1)

C:Species: human herpesvirus 7
A:Virus: strain J1
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T41930
R:Nicholas, J.

Submitted to the EMBL Data Library, December 1995

A:Description: Determination and analysis of the complete nucleotide sequence of human H

A:Reference number: Z22022

A:Accession: T41930

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-806 <NIC>

A:Cross-references: EMBL:U43400; PIDN:AA054690.1

A:Experimental source: strain J1

C:Genetics:

A:Note: U28

Query Match
Best Local Similarity 68.6%; Score 35; DB 2; Length 806;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLMMITOC 9
DB 309 SLMMITOC 317

RESULT 9

S77086
Hypothetical protein sl10737 - *Synechocystis* sp. (strain PCC 6803)

C:Species: *Synechocystis* sp.

A:Virus: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000

C:Accession: S77086

R:Kaneko, T.; Sato, S.; Kocani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

S.

A:Reference number: S74322; NCID:97061201; PMID:8905231

A:Accession: S77086

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-861 <KAN>

A:Cross-references: EMBL:D64005; GB:AB01339; NID:gi1001779; PIDN:BA10778.1; PID:9100662

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Superfamily: *Synechocystis* hypothetical protein sl10737

Query Match
Best Local Similarity 68.6%; Score 35; DB 2; Length 861;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLMMITOC 8
DB 157 SLMMITOC 164

RESULT 10

A86263

F13K23.18 protein - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
C:Accession: A86263
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chen, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luro, J.S.; Malt, R.; Marziani
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86141; NCID:21016719; PMID:11130712
A:Accession: A86263
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-974 <STO>
A:Cross-references: GB:AE005172; NID:96698739; PIDN:AAF78497.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match
Best Local Similarity 68.6%; Score 35; DB 2; Length 974;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LMMITOC 9
DB 557 LMMITOC 564

RESULT 11

F86410
protein F86410 [imported] - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002

C:Accession: F86410

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luro, J.S.; Malt, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A:Reference number: A86141; NCID:21016719; PMID:11130712

A:Accession: F86410

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-996 <STO>

A:Cross-references: GB:AE005172; NID:96560764; PIDN:AAF78497.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

C:Superfamily: protein kinase Xaz1; leucine-rich alpha-2-glycoprotein repeat homology; F

Query Match
Best Local Similarity 68.6%; Score 35; DB 2; Length 996;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LMMITOC 9
DB 209 VMMITOC 215

RESULT 12
T51995
hypothetical protein cur4+ - fission yeast (*Schizosaccharomyces pombe*)
C:Species: *Schizosaccharomyces pombe*
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000

A:Map position: circular chromosome
C:Superfamily: uncharacterized conserved protein

Query Match 70.6%; Score 36; DB 2; Length 198;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLMMIT 7
DB 25 SLMMVT 31

RESULT 3

C97318
uncharacterized conserved membrane protein, probable ABC-type MDR permease CAC3404 [1]
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C/Accession: C97318
R:Noiling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.T.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: C97318
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-247 <KID>
A:Cross-references: GB:AB001437; PIDN:AAK1334.1; PID:815026490; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC3404

Query Match 68.6%; Score 35; DB 2; Length 247;
Best Local Similarity 66.7%; Pred. No. 49;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLMMITOC 9
DB 198 SPLKMITOC 206

RESULT 4
S50336
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 1 - Katharina tunicata mitochondrion
C:Species: mitochondrion Katharina tunicata
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 03-Jun-2002
C/Accession: S50336
R:Boore, J.L.; Brown, W.M.
Genetics 138, 423-443, 1994
A:Title: Complete DNA sequence of the mitochondrial genome of the black chiton, Katharina
A:Reference number: S50327; MUID:95129806; PMID:7828825
A:Accession: S50336
A:Molecule type: DNA
A:Residues: 1-316 <BOO>
A:Cross-references: EMBL:U09810; NID:G557273; PIDN:AA48370.1; PID:G557282
C:Genetics:
A:Gene: NDI

A:Genome: mitochondrion
A:Genetic code: SGC4
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 1
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 68.6%; Score 35; DB 2; Length 316;
Best Local Similarity 75.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 LLMWITOC 9
DB 184 LLMWITVC 191

RESULT 5
T25899

hypothetical protein T20F7.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C/Accession: T25899
R:Miller, N.; Gattung, S.
Submitted to the EMBL Data Library, April 1997
A:Description: The sequence of C. elegans cosmid T20F7.
A:Reference number: Z20107
A:Accession: T25899
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-478 <ML>
A:Cross-references: EMBL:U97550; PIDN:AA852856.1; GSPDB:GN00028; CESP:T20F7.6
C:Experimental source: strain Bristol N2; clone T20F7
C:Genetics:
A:Gene: CESP:T20F7.6
A:Map position: X
A:Introns: 33/3; 112/2; 144/1; 205/3; 263/2; 300/3; 380/3; 402/2; 451/1

Query Match 68.6%; Score 35; DB 2; Length 478;
Best Local Similarity 71.4%; Pred. No. 88;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LLMWITOC 9
DB 43 LLMWITOC 49

RESULT 6

T09024
proline-rich protein T27B11.90 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 22-Oct-1999
C/Accession: T09024
R:Bayan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancr
Submitted to the Protein Sequence Database, June 1999
A:Reference number: Z16533
A:Accession: T09024
A:Molecule type: DNA
A:Residues: 1-577 <BEV>
A:Cross-references: EMBL:AL078579; GSPDB:GN00062; ATSP:T27B11.90
A:Experimental source: cultivar Columbia; BAC clone T27B11
C:Genetics:
A:Gene: ATSP:T27B11.90
A:Map position: 4
A:Introns: 26/1; 117/1; 338/1; 411/3; 430/2; 498/2

Query Match 68.6%; Score 35; DB 2; Length 577;
Best Local Similarity 62.5%; Pred. No. 16+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LLMWITOC 9
DB 405 ILSWITOC 412

RESULT 7

T39684
hypothetical protein SPBC1778.02 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C/Accession: T39684
R:Oliver, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
Submitted to the EMBL Data Library, March 1998
A:Reference number: Z21869
A:Accession: T39684
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-612 <OLI>
A:Cross-references: EMBL:AL049489; PIDN:CA839797.1; GSPDB:GN00067; SPDB:SPBC1778.02
A:Experimental source: strain 972h-; cosmid cl778
C:Genetics:
A:Gene: SPDB:SPBC1778.02

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 17:17:35 ; Search time 10.8621 Seconds

(without alignments)
79.683 Million cell updates/sec

Title: US-10-023-182-5

Perfect score: 51

Sequence: 1 SLMWITQC 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.76.*
2: p1r1.*
3: p1r2.*
4: p1r3.*
5: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	70.6	196	2 AE2876	conserved hypotet
2	36	70.6	198	2 G97652	hypothetical prote
3	35	68.6	247	2 C97318	uncharacterized
4	35	68.6	316	2 S50336	NADH2 dehydrogenas
5	35	68.6	478	2 T25899	hypothetical prote
6	35	68.6	577	2 T09024	proline-rich prote
7	35	68.6	612	2 T39684	hypothetical prote
8	35	68.6	806	2 T41930	ribonucleotide red
9	35	68.6	861	2 S77086	hypothetical prote
10	35	68.6	974	2 A86263	protein F3M18.12 l
11	35	68.6	996	2 T51995	hypothetical prote
12	35	68.6	1458	2 T39266	probable cell cycl
13	35	68.6	1458	2 T39266	ubiquitin-protein
14	35	68.6	1950	2 S12332	genome polyprotein
15	35	68.6	2103	1 UQ1621	probable membrane
16	34	66.7	122	2 S54492	FP26 protein - fow
17	34	66.7	219	2 D35216	coxi intron protei
18	34	66.7	234	2 T11914	probable endoxylor
19	34	66.7	299	2 D84519	hypothetical prote
20	34	66.7	334	2 H69148	hypothetical prote
21	34	66.7	335	2 T03127	hypothetical prote
22	34	66.7	513	2 A37219	mannan endo-1,4-be
23	34	66.7	560	2 T19622	hypothetical prote
24	34	66.7	669	2 A35103	cell surface recep
25	34	66.7	1009	2 S28857	glutamate receptor
26	34	66.7	1009	2 JH0266	glutamate receptor
27	34	66.7	1224	2 T19749	hypothetical prote
28	34	66.7	1253	2 F86436	hypothetical prote
29	33	64.7	113	2 I38320	T-cell receptor be

30	33	64.7	119	2 S08077	Ig kappa chain pre
31	33	64.7	149	2 S74366	hypothetical prote
32	33	64.7	237	2 S35058	GTP-binding protei
33	33	64.7	255	2 A40969	foliate-binding pro
34	33	64.7	274	2 B84930	H+-transporting tw
35	33	64.7	337	2 T12973	protein kinase hom
36	33	64.7	479	2 F84179	hypothetical prote
37	33	64.7	625	2 T24823	hypothetical prote
38	33	64.7	793	2 C83260	hypothetical prote
39	33	64.7	1524	2 T30337	polyprotein - Afri
40	32	62.7	132	1 RMM512	T-cell receptor ga
41	32	62.7	135	1 RMM5V1	T-cell receptor ga
42	32	62.7	137	1 E26420	T-cell receptor ga
43	32	62.7	138	2 G27639	T-cell receptor ga
44	32	62.7	138	2 A27639	T-cell receptor ga
45	32	62.7	138	2 D27639	T-cell receptor ga

ALIGNMENTS

RESULT 1
conserved hypothetical protein Atu2439 [imported] - Agrobacterium tumefaciens (strain C
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Dec-2002
C/Accession: AE2876
R.Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,
erage, G.; Gillet, W.; Grant, C.; Gunttner, D.; Kutyavlin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: AE2876
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-196 <KOR>
A/Cross-references: GB:AE008688; PIDN:AL43427.1; PID:G17740928; GSPDB:GN00186
A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Map position: circular chromosome
A/Gene: Atu2439
C/Superfamily: uncharacterized conserved protein
Query Match
Best Local Similarity 70.6%; Score 36; DB 2; Length 196;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 SLMWIT 7
DB 23 SLMWIT 29
RESULT 2
hypothetical protein AGR_C_4424 [imported] - Agrobacterium tumefaciens (strain C58, Cer
G97652
C/Species: Agrobacterium tumefaciens
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Dec-2002
C/Accession: G97652
R.Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quirello, B.; Goldman
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughly, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu
A/Reference number: A87359; MUID:21608551; PMID:11743194
A/Accession: G97652
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-198 <KUR>
A/Cross-references: GB:AE007869; PIDN:AAK88176.1; PID:G15157620; GSPDB:GN00169
C/Genetics:
A/Gene: AGR_C_4424

```

MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/751.798
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/062.422
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/725.182
FILING DATE: October 3, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. US20020010321A1man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5466.3
TELEPHONE: (212) 318-3168
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 180
TYPE: amino acid
TOPOLOGY: linear
US-09-751-798-8

```

```

Query Match      100.0%; Score 61; DB 9; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.056;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 SLWMITQCF 11
        |||||
        157 SLWMITQCF 167

```

```

RESULT 13
US-09-849-602-30
Sequence 30, Application US/09849602
Publication No. US2003016583A1
GENERAL INFORMATION:
APPLICANT: Scanlan, Matthew J.
APPLICANT: Old, Lloyd J.
APPLICANT: Stockert, Elisabeth
APPLICANT: Chen, Yao-Tseng
TITLE OF INVENTION: Colon Cancer Antigen Panel
FILE REFERENCE: I0461/7105(JRV)
CURRENT APPLICATION NUMBER: US/09/849,602
CURRENT FILING DATE: 2001-05-04
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn version 3.0
SEQ ID NO 30
LENGTH: 180
TYPE: PRT
ORGANISM: Homo sapiens
US-09-849-602-30

```

```

Query Match      100.0%; Score 61; DB 12; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.056;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 SLWMITQCF 11
        |||||
        157 SLWMITQCF 167

```

```

RESULT 14
US-10-026-066-3
Sequence 3, Application US/10026066
Publication No. US20030215425A1
GENERAL INFORMATION:
APPLICANT: Simard, John J. L.
APPLICANT: Diamond, David C.

```

```

TITLE OF INVENTION: EPITOPE SYNCHRONIZATION IN ANTIGEN
FILE REFERENCE: CITIMM.21CPI
CURRENT APPLICATION NUMBER: US/10/026.066
CURRENT FILING DATE: 2001-12-07
PRIOR APPLICATION NUMBER: 09/561,074
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 09/560,465
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 09/561,572
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 09/561,571
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: PCT/US01/13806
PRIOR FILING DATE: 2001-04-27
NUMBER OF SEQ ID NOS: 89
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 180
TYPE: PRT
ORGANISM: Homo sapiens
US-10-026-066-3

```

```

Query Match      100.0%; Score 61; DB 12; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.056;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 SLWMITQCF 11
        |||||
        157 SLWMITQCF 167

```

```

RESULT 15
US-10-117-937-74
Sequence 74, Application US/10117937
Publication No. US20030220239A1
GENERAL INFORMATION:
APPLICANT: CTL IMUNO THERAPIES CORP.
APPLICANT: SIMARD, John, J. L.
APPLICANT: DIAMOND, David, C.
APPLICANT: LIU, Liping
APPLICANT: XIE, Zhidong
TITLE OF INVENTION: EPITOPE SEQUENCES
FILE REFERENCE: CITIMM.027A
CURRENT APPLICATION NUMBER: US/10/117,937
CURRENT FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: US 60/282,211
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: US 60/337,017
PRIOR FILING DATE: 2001-11-07
PRIOR APPLICATION NUMBER: US 60/363,210
PRIOR FILING DATE: 2002-03-07
NUMBER OF SEQ ID NOS: 602
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 74
LENGTH: 180
TYPE: PRT
ORGANISM: Homo sapiens
US-10-117-937-74

```

```

Query Match      100.0%; Score 61; DB 12; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.056;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 SLWMITQCF 11
        |||||
        157 SLWMITQCF 167

```

```

Search completed: February 10, 2004, 17:32:13
Job time : 31.7241 secs

```

FILE REFERENCE: C1005/7012/KA/ERG
CURRENT APPLICATION NUMBER: US/10/161.097
CURRENT FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: US/09/574,749
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: US 60/107,972
PRIOR FILING DATE: 1998-11-12
PRIOR APPLICATION NUMBER: PCT/US99/26795
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: US 09/524,749
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 58
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 33
LENGTH: 11
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Homo Sapiens source
US-10-161-097-33

Query Match 100.0%; Score 61; DB 15; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0053;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLWMITQCFL 11
Db 1 SLWMITQCFL 11

RESULT 10
US-10-117-937-255
Sequence 255, Application US/10117937
Publication No. US2003022039A1
GENERAL INFORMATION:
APPLICANT: CTL IMMUNO THERAPIES CORP.
APPLICANT: SIMARD, John, J.L.
APPLICANT: DIAMOND, David, C.
APPLICANT: LIT, Liping
APPLICANT: XIE, Zhidong
TITLE OF INVENTION: EPITOPE SEQUENCES
FILE REFERENCE: CTLIM.027A
CURRENT APPLICATION NUMBER: US/10/117,937
CURRENT FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: US 60/282,211
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: US 60/337,017
PRIOR FILING DATE: 2001-11-07
PRIOR APPLICATION NUMBER: US 60/363,210
PRIOR FILING DATE: 2002-03-07
NUMBER OF SEQ ID NOS: 602
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 255
LENGTH: 28
TYPE: PRT
ORGANISM: Homo sapiens
US-10-117-937-255

Query Match 100.0%; Score 61; DB 12; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLWMITQCFL 11
Db 8 SLWMITQCFL 18

RESULT 11
US-10-295-027-388
Sequence 388, Application US/10295027
Publication No. US2003023250A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel

APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezi, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Ros Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 388
LENGTH: 135
TYPE: PRT
ORGANISM: Homo sapiens
US-10-295-027-388

Query Match 100.0%; Score 61; DB 12; Length 135;
Best Local Similarity 100.0%; Pred. No. 0.044;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLWMITQCFL 11
Db 112 SLWMITQCFL 122

RESULT 12
US-09-751-798-8
Sequence 8, Application US/09751798
Patent No. US20020010321A1
GENERAL INFORMATION:
APPLICANT: Stockert, Elisabeth; Jager, Elke;
APPLICANT: Chen, Yao-tseng; Scanlan, Matthew;
APPLICANT: Knuth, Alexander; Old, Lloyd U.
TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer
TITLE OF INVENTION: Associated Proteins, Uses Thereof,
TITLE OF INVENTION: Truncated Forms of NY-ESO-1, and HLA
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski, L.L.P.
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10103
COMPUTER READABLE FORM:

LENGTH: 11
TYPE: PRT
ORGANISM: Homo sapiens
US-10-177-277-130

Query Match 100.0%; Score 61; DB 12; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLMMWITQCF 11
DB 1 SLMMWITQCF 11

RESULT 6
US-10-164-078A-28
Sequence 28, Application US/10164078A
Publication No. US20030228325A1
GENERAL INFORMATION:
APPLICANT: Bilsborough, Janine
APPLICANT: Schultz, Erwin
APPLICANT: Panichelli, Christophe
APPLICANT: Van der Bruggen, Pierre
APPLICANT: Boon, Thierry
TITLE OF INVENTION: Isolated Peptides Which Bind to HLA-B*8 Molecules And Uses Thereof
FILE REFERENCE: LUD-5756
CURRENT APPLICATION NUMBER: US/10/164,078A
CURRENT FILING DATE: 2002-06-05
NUMBER OF SEQ ID NOS: 35
SEQ ID NO 28
LENGTH: 11
TYPE: PRT
ORGANISM: H. sapiens
FEATURE:
US-10-164-078A-28

Query Match 100.0%; Score 61; DB 12; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLMMWITQCF 11
DB 1 SLMMWITQCF 11

RESULT 7
US-10-164-121A-29
Sequence 29, Application US/10164121A
Publication No. US20030228308A1
GENERAL INFORMATION:
APPLICANT: Zhang, Yi
APPLICANT: Boon, Thierry
APPLICANT: Van der Bruggen, Pierre
APPLICANT: Traversari, Carla
TITLE OF INVENTION: Isolated Peptides Which Bind to HLA-Cw6 Molecules And Uses Thereof
FILE REFERENCE: LUD-5771
CURRENT APPLICATION NUMBER: US/10/164,121A
CURRENT FILING DATE: 2002-08-26
NUMBER OF SEQ ID NOS: 36
SEQ ID NO 29
LENGTH: 11
TYPE: PRT
ORGANISM: H. sapiens
FEATURE:
US-10-164-121A-29

Query Match 100.0%; Score 61; DB 12; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLMMWITQCF 11
DB 1 SLMMWITQCF 11

RESULT 8
US-10-023-182-4
Sequence 4, Application US/10023182
Publication No. US20020164655A1
GENERAL INFORMATION:
APPLICANT: Stockert, Elisabeth; Jager, Elke;
Chen, Yao-tsen; Scanlan, Matthew;
Knuth, Alexander; Old, Lloyd J.

TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer Associated Proteins, Uses Thereof, Truncated Forms of NY-ESO-1, and HLA Binding Peptides Derived Therefrom

NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski, L.L.P.
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10103

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/023,182
FILING DATE: 17-Dec-2001

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/751,798
FILING DATE: December 29, 2000
APPLICATION NUMBER: 09/062,422

FILING DATE: <Unknown>
APPLICATION NUMBER: 08/725,182
FILING DATE: October 3, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. US20020164655A1man D.
REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: LUD 5466.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3168
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-10-023-182-4
Query Match 100.0%; Score 61; DB 14; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLMMWITQCF 11
DB 1 SLMMWITQCF 11

RESULT 9
US-10-161-097-33
Sequence 33, Application US/10161097
Publication No. US20030096404A1
GENERAL INFORMATION:
APPLICANT: ROSENZWEIG, Michael
APPLICANT: PYKEIT, Mark J.
APPLICANT: SCADENSKY, David T.
APPLICANT: POZNANSKY, Mark C.
TITLE OF INVENTION: LIVEROID TISSUE-SPECIFIC CELL PRODUCTION
TITLE OF INVENTION: FROM HEMATOPOIETIC PROGENITOR CELLS IN THREE-DIMENSIONAL

Fri Feb 13 14:37:38 2004

US-10-023-182-4.rapb

Page 2

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3168
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-751-798-4

Query Match 100.0%; Score 61; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLWMWITQCFL 11
DB 1 SLWMWITQCFL 11

RESULT 2
US-09-766-889A-50
Sequence 50, Application US/09766889A
Patent No. US2002016454A1
GENERAL INFORMATION:
APPLICANT: Luiten, Rosalie
APPLICANT: Boon-Falleur, Thierry
APPLICANT: van der Bruggen, Pierre
APPLICANT: Stroobant, Vincent
APPLICANT: Demotte, Nathalie
APPLICANT: Schulte, Edwin
TITLE OF INVENTION: MAGE ANTIGENIC PEPTIDES WHICH BIND HLA-B35 AND HLA-B44
FILE REFERENCE: L0461/7104
CURRENT APPLICATION NUMBER: US/09/766,889A
CURRENT FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: US 60/177,242
PRIOR FILING DATE: 2000-01-20
PRIOR APPLICATION NUMBER: US 60/243,212
PRIOR FILING DATE: 2000-10-25
NUMBER OF SEQ ID NOS: 59
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 50
LENGTH: 11
TYPE: PRT
ORGANISM: Homo sapiens
US-09-766-889A-50

Query Match 100.0%; Score 61; DB 10; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLWMWITQCFL 11
DB 1 SLWMWITQCFL 11

RESULT 3
US-10-170-832-79
Sequence 79, Application US/10170832
Publication No. US20030170792A1
GENERAL INFORMATION:
APPLICANT: Vantomme, Valerie
APPLICANT: Chaub, Pascal
APPLICANT: Stroobant, Vincent
APPLICANT: Boon-Falleur, Thierry
APPLICANT: van der Bruggen, Pierre
APPLICANT: Thielemans, Kees
APPLICANT: Corthals, Jürgen
TITLE OF INVENTION: MAGE-3 PEPTIDES PRESENTED BY HLA CLASS II MOLECULES
FILE REFERENCE: L0461/7052
CURRENT APPLICATION NUMBER: US/10/170,832
CURRENT FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: US/09/166,448
PRIOR FILING DATE: 1998-10-05

NUMBER OF SEQ ID NOS: 81
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 79
LENGTH: 11
TYPE: PRT
ORGANISM: Homo sapiens
US-10-170-832-79

Query Match 100.0%; Score 61; DB 12; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLWMWITQCFL 11
DB 1 SLWMWITQCFL 11

RESULT 4
US-10-364-614-9
Sequence 9, Application US/10364614
Publication No. US20030175250A1
GENERAL INFORMATION:
APPLICANT: JAGER, Elke
APPLICANT: KNUTH, Alexander
APPLICANT: OLD, Lloyd
APPLICANT: Gnifatic, Sacha
TITLE OF INVENTION: ISOLATED PEPTIDES WHICH BIND TO HLA MOLECULES AND USES THEREOF
FILE REFERENCE: LUD 5726.1 CIP
CURRENT APPLICATION NUMBER: US/10/364,614
CURRENT FILING DATE: 2003-02-24
PRIOR APPLICATION NUMBER: US 60/355,828
PRIOR FILING DATE: 2002-02-13
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn version 3.2
SEQ ID NO 9
LENGTH: 11
TYPE: PRT
ORGANISM: Homo sapiens
US-10-364-614-9

Query Match 100.0%; Score 61; DB 12; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLWMWITQCFL 11
DB 1 SLWMWITQCFL 11

RESULT 5
US-10-177-277-130
Sequence 130, Application US/10177277
Publication No. US2003018584A1
GENERAL INFORMATION:
APPLICANT: Tureci, Ozlem
APPLICANT: Sahin, Ugur
APPLICANT: Pfeundschnuh, Michael
APPLICANT: Ramensee, Hans Georg
APPLICANT: Stevanovic, Stefan
TITLE OF INVENTION: Method for Determining Presence of Cancer in a Sample By Determining Expression of an SSX Gene, Peptides Derived From Said SSX Gene
FILE REFERENCE: LUD 5556.1
CURRENT APPLICATION NUMBER: US/10/177,277
CURRENT FILING DATE: 2002-06-21
PRIOR APPLICATION NUMBER: US/09/344,040
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 09/105,839
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: US 08/851,130
PRIOR FILING DATE: 1997-05-05
NUMBER OF SEQ ID NOS: 132
SEQ ID NO 130

Fri Feb 13 14:37:38 2004

us-10-023-182-4.rapb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 17:19:51 ; Search time 30.7241 Seconds
(without alignments)
74.964 Million cell updates/sec

Title: US-10-023-182-4
Perfect score: 61
Sequence: 1 SLMMWITQCFLL 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	11	9	US-09-751-798-4
2	61	100.0	11	10	US-09-766-889A-50
3	61	100.0	11	12	US-10-170-832-79
4	61	100.0	11	12	US-10-364-614-9
5	61	100.0	11	12	US-10-177-277-130
6	61	100.0	11	12	US-10-164-078A-38
7	61	100.0	11	12	US-10-164-121A-29
8	61	100.0	11	14	US-10-023-182-4
9	61	100.0	11	15	US-10-161-097-33
10	61	100.0	28	12	US-10-117-937-255
11	61	100.0	135	12	US-10-295-027-388
12	61	100.0	180	9	US-09-751-798-8
13	61	100.0	180	12	US-09-849-602-30
14	61	100.0	180	12	US-10-026-065-3
15	61	100.0	180	12	US-10-117-937-74

16	61	100.0	180	12	US-10-117-937-75	Sequence 75, Appl
17	61	100.0	180	12	US-10-295-027-386	Sequence 386, App
18	61	100.0	180	14	US-10-023-182-8	Sequence 8, Appl
19	61	100.0	180	15	US-10-146-473-69	Sequence 69, Appl
20	61	100.0	180	15	US-10-207-655-71	Sequence 71, Appl
21	61	100.0	397	9	US-09-821-883-27	Sequence 27, Appl
22	54	88.5	180	12	US-10-364-614-14	Sequence 14, Appl
23	53	86.9	9	12	US-09-833-039-123	Sequence 123, App
24	53	86.9	9	12	US-10-177-277-123	Sequence 123, App
25	51	83.6	9	12	US-09-751-798-5	Sequence 51, Appl
26	51	83.6	9	10	US-09-766-889A-51	Sequence 51, Appl
27	51	83.6	9	12	US-10-170-832-80	Sequence 80, Appl
28	51	83.6	9	12	US-10-364-614-10	Sequence 10, Appl
29	51	83.6	9	12	US-10-177-277-131	Sequence 131, App
30	51	83.6	9	12	US-10-117-937-137	Sequence 137, App
31	51	83.6	9	12	US-10-164-078A-14	Sequence 14, Appl
32	51	83.6	9	12	US-10-164-121A-15	Sequence 15, Appl
33	51	83.6	9	14	US-10-023-182-5	Sequence 5, Appl
34	51	83.6	9	15	US-10-274-017-1	Sequence 1, Appl
35	51	83.6	9	15	US-10-080-013-24	Sequence 24, Appl
36	51	83.6	9	15	US-10-161-097-34	Sequence 16, App
37	44	72.1	9	12	US-10-117-937-136	Sequence 495, App
38	44	72.1	20	12	US-10-313-986-495	Sequence 265, App
39	44	72.1	151	12	US-10-264-049-2659	Sequence 2047, Ap
40	44	72.1	593	12	US-10-094-749-2047	Sequence 25, Appl
41	42	68.9	9	15	US-10-080-013-25	Sequence 4233, Ap
42	38	62.3	462	12	US-10-369-493-4233	Sequence 6, Appl
43	37	60.7	9	9	US-09-751-798-6	Sequence 52, Appl
44	37	60.7	9	10	US-09-766-889A-52	Sequence 81, Appl
45	37	60.7	9	12	US-10-170-832-81	

ALIGNMENTS

RESULT 1
US-09-751-798-4
; Sequence 4, Application US/09751798
; Patent No. US20020010321A1
; GENERAL INFORMATION:
; APPLICANT: Stockert, Elisabeth; Jager, Elke;
; APPLICANT: Chen, Yao-tseng; Scanlan, Matthew;
; APPLICANT: Knuth, Alexander; Old, Lloyd J.
; TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer
; TITLE OF INVENTION: Associated Proteins, Uses Thereof,
; TITLE OF INVENTION: Truncated Forms of NY-ESO-1, and HLA
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski, L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/751,798
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/062,422
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/725,182
; FILING DATE: October 3, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. US20020010321A1man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: UDD. 5466.3

DR Pfam; PF03010; GP4; 1.
SQ SEQUENCE 152 AA; 17082 MW; CP18A8AA1B9D756B CRC64;
Query Match
Best Local Similarity 62.3%; Score 38; DB 12; Length 152;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 2 LLMWITOCPL 11
| : | : | : |
Db 143 LVRWVQCCYL 152

RESULT 14
Q9QKH1 PRELIMINARY; PRT; 152 AA.
AC Q9QKH1;
DT 01-MAY-2000 (Tremblrel.13, Created)
DT 01-MAY-2000 (Tremblrel.13, Last sequence update)
DT 01-DEC-2001 (Tremblrel.19, Last annotation update)
DE ORF4.
OS Equine arteritis virus (EAV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Arteriviridae; Arterivirus.
OX NCBI_TaxID=11047;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A1;
RX MEDLINE=9394666; PubMed=10466790;
RA Balasuriya U.B., Hedges J.F., Nadler S.A., McColium W.H.,
RA Timoney P.J., MacLachlan N.J.;
RT "Genetic stability of equine arteritis virus during horizontal and
RT vertical transmission in an outbreak of equine viral arteritis.";
RL J. Gen. Virol. 80:1949-1958 (1999).
DR EMBL; AF18769; A054683.1; -.
DR InterPro; IPR004257; GP4.
DR Pfam; PF03010; GP4; 1.
SQ SEQUENCE 152 AA; 16995 MW; 5F8FE7815B70835F CRC64;
Query Match
Best Local Similarity 62.3%; Score 38; DB 12; Length 152;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 2 LLMWITOCPL 11
| : | : | : |
Db 143 LVRWVQCCYL 152

RESULT 15
Q9QKF2 PRELIMINARY; PRT; 152 AA.
AC Q9QKF2;
DT 01-MAY-2000 (Tremblrel.13, Created)
DT 01-MAY-2000 (Tremblrel.13, Last sequence update)
DT 01-DEC-2001 (Tremblrel.19, Last annotation update)
DE ORF4.
OS Equine arteritis virus (EAV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Arteriviridae; Arterivirus.
OX NCBI_TaxID=11047;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RO;
RX MEDLINE=9394666; PubMed=10466790;
RA Balasuriya U.B., Hedges J.F., Nadler S.A., McColium W.H.,
RA Timoney P.J., MacLachlan N.J.;
RT "Genetic stability of equine arteritis virus during horizontal and
RT vertical transmission in an outbreak of equine viral arteritis.";
RL J. Gen. Virol. 80:1949-1958 (1999).
DR EMBL; AF18782; A054761.1; -.
DR InterPro; IPR004257; GP4.
DR Pfam; PF03010; GP4; 1.
SQ SEQUENCE 152 AA; 17138 MW; 0E5AF3C417BC256F CRC64;

Query Match
Best Local Similarity 62.3%; Score 38; DB 12; Length 152;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 2 LLMWITOCPL 11
| : | : | : |
Db 143 LVRWVQCCYL 152
Search completed: February 10, 2004, 17:19:43
Job time : 34.6621 secs

Db 6 LTMWVACFL 15

RESULT 10

ID 080372 PRELIMINARY; PRT; 107 AA.

AC 080372;

DT 01-NOV-1998 (TREMBLrel. 08, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE SAMIPB (Fragment).

CN SAMIPB

OS Aster tripolium (Sea aster).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; campanulids; Asterales; Asteraceae; Asteroideae; Astereae;

OC Aster.

CX NCBI_TaxID=74787;

RN [1]

RP SEQUENCE FROM N.A.

RA Uno Y., Urao T., Shinozaki K.Y., Kanechi M., Inagaki N., Maekawa S.,

RA Shinozaki K.;

RT "Early Salt-Stress Effects on Expression of Genes for Aquaporin

Homologues in the Halophyte Sea Aster (Aster tripolium L.).";

RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.

CC -1 SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).

CC -1 SIMILARITY: BELONGS TO THE MIP/AQUAPORIN FAMILY (TC 1.A.8).

DR EMBL; AB012268; BA31516.1; -.

DR HSSP; P47865; 144N.

DR InterPro; IPR00425; MIP_family.

DR Pfam; PF00230; MIP; 1.

DR PRINTS; PR00783; MINTRINSICP.

DR ProDom; PD000295; MIP_family; 1.

DR PROSITE; PS00221; MIP; 1.

FM Transmembrane; Transport.

FT NON_TER

SQ SEQUENCE 107 AA; 11070 MW; F39B5840CEFC7332 CRC64;

QY 1 SLWMITQC 9

Db 23 SILWIAQC 31

RESULT 11

ID 09PX34 PRELIMINARY; PRT; 152 AA.

AC 09PX34;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE ORF4.

OS Equine arteritis virus (EAV).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;

OC Arteriviridae; Arterivirus.

CX NCBI_TaxID=11047;

RN [1]

RP SEQUENCE FROM N.A.

RA Balasuriya U.B.R., Hedges J.F., Nadler S.A., McCollum W.H.,

RA Timoney P.J., MacLachlan N.J.;

RT "Genetic Stability of Equine Arteritis Virus during Horizontal and

Vertical Transmission in an Outbreak of Equine Viral Arteritis.";

RL J. Gen. Virol. 80:1949-1958(1999).

DR EMBL; AF118776; AAD54725.1; -.

DR EMBL; AF118770; AAD54689.1; -.

DR EMBL; AF118771; AAD54695.1; -.

DR EMBL; AF118772; AAD54701.1; -.

DR InterPro; IPR004257; GP4.

DR Pfam; PF03010; GP4; 1.

SQ SEQUENCE 152 AA; 17124 MW; 5FEFF3D15420985P CRC64;

QY 2 LLMWITQCFL 11

Db 143 LVRWVQCFL 152

RESULT 12

ID 09PX35 PRELIMINARY; PRT; 152 AA.

AC 09PX35;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE ORF4.

OS Equine arteritis virus (EAV).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;

OC Arteriviridae; Arterivirus.

CX NCBI_TaxID=11047;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN=RT-PA96, RI, and PI.

RA MEDLINE=99394666; PubMed=10466790;

RA Balasuriya U.B.R., Hedges J.F., Nadler S.A., McCollum W.H.,

RA Timoney P.J., MacLachlan N.J.;

RT "Genetic Stability of Equine Arteritis Virus during Horizontal and

Vertical Transmission in an Outbreak of Equine Viral Arteritis.";

RL J. Gen. Virol. 80:1949-1958(1999).

DR EMBL; AF118781; AAD54725.1; -.

DR EMBL; AF118773; AAD54707.1; -.

DR EMBL; AF118775; AAD54719.1; -.

DR InterPro; IPR004257; GP4.

DR Pfam; PF03010; GP4; 1.

SQ SEQUENCE 152 AA; 17150 MW; 5E47F3CD85E193B4 CRC64;

QY 2 LLMWITQCFL 11

Db 143 LVRWVQCFL 152

RESULT 13

ID 09OKG6 PRELIMINARY; PRT; 152 AA.

AC 09OKG6;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE ORF4.

OS Equine arteritis virus (EAV).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;

OC Arteriviridae; Arterivirus.

CX NCBI_TaxID=11047;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN=R2;

RA MEDLINE=99394666; PubMed=10466790;

RA Balasuriya U.B.R., Hedges J.F., Nadler S.A., McCollum W.H.,

RA Timoney P.J., MacLachlan N.J.;

RT "Genetic stability of equine arteritis virus during horizontal and

vertical transmission in an outbreak of equine viral arteritis.";

RL J. Gen. Virol. 80:1949-1958(1999).

DR EMBL; AF118774; AAD54713.1; -.

DR InterPro; IPR004257; GP4.

RT to the human gastrointestinal tract.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
 DR EMBL; AB014751; AAN25052.1; -
 KW Antinocetyl-TRNA synthetase; Complete proteome.
 SQ SEQUENCE 379 AA; 42168 MW; DEAC7D620B757C6B CRC64;

Query Match 65.6%; Score 40; DB 16; Length 379;
 Best Local Similarity 63.6%; Pred. No. 40;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SLMMITQCF 11
 DB 256 ALQWIRQCL 266

RESULT 6
 Q962M9 PRELIMINARY; PRT; 273 AA.
 AC Q962M9;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DE PVLH14025 P.
 GN PVLH14025G.
 OS Plasmodium vivax.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
 OX NCBI_TaxID=5855;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tchavatchitch M., Fischer K., Huestis R., Saul A.;
 RT "The sequence of 200 kb portion of a Plasmodium vivax chromosome
 RT reveals a high degree of conservation with P. falciparum chromosome
 RT 3.";
 RL Mol. Biochem. Parasitol. 0:0-0(2001).
 DR EMBL; AY003872; AAF99451.1; -
 SQ SEQUENCE 273 AA; 32028 MW; 73F0E166635A30C CRC64;

Query Match 63.9%; Score 39; DB 5; Length 273;
 Best Local Similarity 60.0%; Pred. No. 45;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SLMMITQCF 10
 DB 26 TLIMITQCF 35

RESULT 7
 Q8GUB0 PRELIMINARY; PRT; 353 AA.
 AC Q8GUB0;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DE Putative vacuolar ATPase subunit d (EC 3.6.1.34).
 GN VBA-D.
 OS Mesembryanthemum crystallinum (Common ice plant).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllales; Caryophyllales; Alzooceae; Mesembryanthemum.
 OX NCBI_TaxID=3544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RA Kluge C., Tavaoli N., Goldack D., Dietz K.J.;
 RT "Cloning and expression analysis of all cDNAs coding for the vacuolar
 RT ATPase in M. crystallinum.";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ39342; CAD27914.1; -
 KW Hydrolyase.
 SQ SEQUENCE 353 AA; 40291 MW; B483F89643E056A2 CRC64;

Query Match 63.9%; Score 39; DB 10; Length 353;
 Best Local Similarity 85.7%; Pred. No. 57;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 3 LMMITQC 9
 DB 330 LMMISQC 336

RESULT 8
 Q8COTO PRELIMINARY; PRT; 401 AA.
 ID Q8COTO;
 AC Q8COTO;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DE Hypothetical type-1 copper.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the Riken Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 DR EMBL; AK029920; BAC26677.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 401 AA; 45562 MW; F34F237653DFCF9 CRC64;

Query Match 63.1%; Score 38.5; DB 11; Length 401;
 Best Local Similarity 61.5%; Pred. No. 77;
 Matches 8; Conservative 1; Mismatches 1; Indels 3; Gaps 1;
 QY 1 SLMMIT---OCF 10
 DB 316 SLMMVITGACF 328

RESULT 9
 Q46517 PRELIMINARY; PRT; 65 AA.
 ID Q46517;
 AC Q46517;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DE Desulfurococcus mobilis.
 ORF 65 (Fragment).
 OS Desulfurococcus mobilis.
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;
 OC Desulfurococcaceae; Desulfurococcus.
 OX NCBI_TaxID=2274;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 2126;
 RX MEDLINE=95206243; PubMed=7898436;
 RA Ceccarelli E., Bocchetta M., Creti R., Sanangelantoni A.M., Tiboni O.,
 RA Camarero P.;
 RT "Chromosomal organization and nucleotide sequence of the genes for
 RT elongation factors EF- α and EF- β and ribosomal proteins S7 and
 RT S10 of the hyperthermophilic archaeum Desulfurococcus mobilis.";
 RL Mol. Gen. Genet. 246:687-696(1995).
 DR EMBL; X73582; CAA51989.1; -
 FT NON TER 65
 SQ SEQUENCE 65 AA; 7597 MW; B51DEF4F5C84E0F8 CRC64;

Query Match 62.3%; Score 38; DB 1; Length 65;
 Best Local Similarity 60.0%; Pred. No. 20;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 2 LMMITQCF 11

Db 157 SLMMITOCFL 167

RESULT 2

Q9JUK3 PRELIMINARY; PRT; 151 AA.
 AC Q9JUK3;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE D3J10U6.1 (Novel protein) (Fragment).
 GN D3J10U6.1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Blakey S.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL035593; CAB56526.1; -.
 FT NON TER 1
 SO SEQUENCE 151 AA; 17990 MW; 9895001B4CF4B6DE CRC64;

Query Match 72.1%; Score 44; DB 4; Length 151;
 Best Local Similarity 87.5%; Pred. No. 3.8;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LMMITOCF 10
 Db 52 LMMITOCF 59

RESULT 3

Q96NH3 PRELIMINARY; PRT; 593 AA.
 AC Q96NH3;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein FLJ30899.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Brain;
 RA Nishi T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H.,
 RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
 RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
 RA Kikuchi H., Kanda K., Matsuo K., Nakamura Y., Sekine M.,
 RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
 RA Sugano S., Nagahara K., Masuno Y., Nagai K., Isogai T.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK055461; BAB70925.1; -.
 KW Hypothetical protein.
 SO SEQUENCE 593 AA; 66636 MW; B3295013BB225BC CRC64;

Query Match 72.1%; Score 44; DB 4; Length 593;
 Best Local Similarity 87.5%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LMMITOCF 10
 Db 494 LMMITOCF 501

RESULT 4
 Q8UCP4 PRELIMINARY; PRT; 198 AA.

AC Q8UCP4;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein Atu2439.
 GN Atu2439 OR AGR C.4424.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Sebubal J.C., Kaul R., Monks D.E., Kiteajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D., Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutyavin I., Levy R., Li M.-J., McClelland E., Palmieri A.,
 RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung W., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 C58.";
 RL Science 294:2317-2323 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Gourello B., Goldman B.S., Cao Y., Askenazi M., Halling L.,
 RA Hounel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Dougherty D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328 (2001).
 DR EMBL; AE009191; AA143427.1; ALT_INIT.
 DR EMBL; AE008157; AA08176.1; -.
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 198 AA; 21154 MW; FA40CE548FD562P CRC64;

Query Match 67.2%; Score 41; DB 16; Length 198;
 Best Local Similarity 70.0%; Pred. No. 16;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SLMMITOCF 10
 Db 25 SLMMITOCF 34

RESULT 5
 Q8G4X3 PRELIMINARY; PRT; 379 AA.
 AC Q8G4X3;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Probable glutamyl-tRNA synthetase.
 GN GLTY2 OR BL1251.
 OS Bifidobacterium longum.
 OC Bacteria; Actinobacteria; Actinobacteriales; Bifidobacteriales;
 OC Bifidobacteriaceae; Bifidobacterium.
 OX NCBI_TaxID=216816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=NC 2705;
 RX MEDLINE=22294977; PubMed=12381787;
 RA Schell M.A., Karmaliantzou M., Snel B., Vilanova D., Berger B.,
 RA Pessi G., Zwaan M.-C., Desiere F., Bork P., Delley M.,
 RA Pridmore R.D., Arigon F.;
 RT "The genome sequence of Bifidobacterium longum reflects its adaptation

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 17:14:10 ; Search time 31.8621 Seconds
(without alignments)
89,090 Million cell updates/sec

Title: US-10-023-182-4

Perfect score: 61

Sequence: 1 SLWMITQCF 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP viirus:*
16: SP bacteriap:*
17: SP archaep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	180	4 Q9Y479	Q9Y479 homo sapien
2	44	72.1	151	4 Q9UKX3	Q9UKX3 homo sapien
3	44	72.1	593	4 Q96NH3	Q96NH3 homo sapien
4	41	67.2	198	16 Q8UCP4	Q8UCP4 agrobacteri
5	40	65.6	379	16 Q8G4X3	Q8G4X3 difidobacte
6	39	63.9	273	5 Q96ZM9	Q96ZM9 plasmodium
7	39	63.9	353	10 Q8GUB0	Q8GUB0 mesembryant
8	38.5	62.3	401	11 Q8COT0	Q8COT0 mus musculu
9	38	62.3	65	1 Q46517	Q46517 desulfuroco
10	38	62.3	107	10 Q80372	Q80372 asper tripo
11	38	62.3	152	12 Q9PX34	Q9PX34 equine arte
12	38	62.3	152	12 Q9PX35	Q9PX35 equine arte
13	38	62.3	152	12 Q9GK6	Q9GK6 equine arte
14	38	62.3	152	12 Q9GKH1	Q9GKH1 equine arte
15	38	62.3	152	12 Q9GKF2	Q9GKF2 equine arte
16	38	62.3	162	16 Q8D2P2	Q8D2P2 wigleswort

17	38	62.3	317	5 Q97336	Q97336 plasmodium
18	38	62.3	408	16 Q9KRM7	Q9KRM7 vibrio chol
19	38	62.3	423	5 Q02168	Q02168 caenorhabd
20	38	62.3	490	11 Q91Z08	Q91Z08 mus musculu
21	38	62.3	532	3 Q8WXY1	Q8WXY1 neurospora
22	38	62.3	814	11 Q8CED5	Q8CED5 mus musculu
23	38	62.3	996	10 Q9SGP2	Q9SGP2 mus musculu
24	38	62.3	2664	6 Q8MN95	Q8MN95 bos taurus
25	37.5	61.5	83	12 Q9WGD6	Q9WGD6 human papil
26	37	60.7	51	16 Q8FE64	Q8FE64 escherichia
27	37	60.7	165	11 Q8R001	Q8R001 mus musculu
28	37	60.7	165	11 Q8BNX7	Q8BNX7 mus musculu
29	37	60.7	159	10 Q9ZMN0	Q9ZMN0 phaseolus a
30	37	60.7	511	5 Q96332	Q96332 heliothis v
31	37	60.7	776	3 Q8X013	Q8X013 neurospora
32	37	60.7	783	3 Q14411	Q14411 candida alb
33	37	60.7	1189	5 Q9VQ78	Q9VQ78 drosophila
34	36.5	59.8	581	17 Q29674	Q29674 archaeoglob
35	36	59.0	70	12 Q90748	Q90748 encephalomy
36	36	59.0	71	12 Q90751	Q90751 encephalomy
37	36	59.0	101	10 Q9AXA1	Q9AXA1 oryza sativ
38	36	59.0	185	10 Q8S0K9	Q8S0K9 oryza sativ
39	36	59.0	193	16 Q8ZL27	Q8ZL27 salmoneilla
40	36	59.0	193	16 Q8Z3P1	Q8Z3P1 salmoneilla
41	36	59.0	209	10 Q8GWC3	Q8GWC3 arabidopsis
42	36	59.0	218	10 Q9LRK4	Q9LRK4 arabidopsis
43	36	59.0	223	10 Q9LRK5	Q9LRK5 arabidopsis
44	36	59.0	245	11 Q8ATP0	Q8ATP0 mus musculu
45	36	59.0	253	5 Q8IGP6	Q8IGP6 drosophila

ALIGNMENTS

RESULT 1

Q9Y479 Q9Y479 PRELIMINARY; PRT; 180 AA.
ID Q9Y479
AC Q9Y479/
DT 01-NOV-1999 (TREMUR1.12, Created)
DT 01-NOV-1999 (TREMUR1.12, Last sequence update)
DT 01-OCT-2002 (TREMUR1.22, Last annotation update)
DE LAGE-15 protein (Cancer/testis antigen 2).
GN LAGE1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=99325550; PubMed=1039963;
RA Aarnoudse C.A., Van den Doel P.B., Heemskerk B., Schrier P.I.;
RT "Interleukin-2-induced, melanoma-specific T cells recognize CAMEL, an
RT unexpected translation product of LAGE-1.";
RL Int. J. Cancer 82:442-448(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Aradhye S., Bardaro T., Galigoczy P., Yamagata T., Esposto T.,
RA Patlan H., Ciccodicola A., Kenwright S., Platzner M., D'Urso M.,
RA Nelson D.L.;
RT "Multiple pathogenic and benign genomic rearrangements occur at a 35-
RT kb duplication involving the NEMO and the LAGE2 genes.";
RH Hum. Mol. Genet. 0:0-0(2001).
DR EMBL; AF012834; CAA0194.1;
DR EMBL; AF277315; AAL27015.1;
SQ SEQUENCE 180 AA; 18236 MW; 9077FAP953543A25 CRC64;

Query Match 100.0%; Score 61; DB 4; Length 180;
Best local similarity 100.0%; Pred. No. 0.0053;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLWMITQCF 11
|||||

```

RN      J. Biol. Chem. 265:4472-4482(1990).
RP      [3]
RX      MEDLINE=89140821; PubMed=2918319;
RA      Iseberg K.E., Meyer G.E.;
RT      "Cloning of a putative neuronal nicotinic acetylcholine receptor
        subunit.";
RL      J. Neurochem. 52:988-991(1989).
CC      -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN
CC      EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND
CC      LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA
CC      MEMBRANE.
CC      -1- SUBUNIT: NEURONAL ACHR SEEMS TO BE COMPOSED OF TWO DIFFERENT TYPES
CC      OF SUBUNITS: ALPHA-2, ALPHA-3, OR ALPHA-4 SUBUNIT CAN BE
CC      COMBINED WITH ALPHA-2, ALPHA-3, OR ALPHA-4 SUBUNITS TO GIVE RISE TO
CC      FUNCTIONAL RECEPTORS.
CC      -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC      -1- ALTERNATIVE PRODUCTS:
CC      Event=Alternative splicing; Named Isoforms=2;
CC      Name=Short;
CC      IsoId=P12392-1; Sequence=Displayed;
CC      Name=Long;
CC      IsoId=P12392-2; Sequence=VSP 000076;
CC      -1- TISSUE SPECIFICITY: IN THE BRAIN, IT IS DETECTED IN THE MEDIAL
CC      HABENULA. IN THE PERIPHERAL NERVOUS SYSTEM, IT IS FOUND AT LEAST
CC      IN THE ADRENAL GLAND.
CC      -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-stb.ch/announce/
CC      or send an email to license@isb-stb.ch).
CC      -----
DR      EMBL; U42976; AAA85212.1; -.
DR      EMBL; M33953; AAA41668.1; JOINED.
DR      EMBL; J05232; AAA41668.1; JOINED.
DR      EMBL; M89871; AAA41668.1; JOINED.
DR      EMBL; M33851; AAA41668.1; JOINED.
DR      EMBL; M89889; AAA41668.1; JOINED.
DR      EMBL; M33952; AAA41668.1; JOINED.
DR      EMBL; X15834; CAA33839.1; -.
DR      PIR; B35721; B35721.
DR      InterPro; IPRO06029; Neu_channel memb.
DR      InterPro; IPRO06202; Neur_chan_lbd.
DR      InterPro; IPRO06201; Neur_channel.
DR      Pfam; PF02931; Neur_chan_LBD; 1.
DR      Pfam; PF02932; Neur_chan_memb; 1.
DR      PRINTS; PR00252; NRIONCHANNEL.
DR      TIGRFAMs; TIGR00860; LIC; 1.
DR      PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW      Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW      Transmembrane; Multigene family; Alternative splicing.
FT      SIGNAL          1    20
FT      CHAIN           21   495
FT      DOMAIN         21   230
FT      TRANSMEM     231   255
FT      TRANSMEM     263   281
FT      TRANSMEM     297   318
FT      DOMAIN       319   463
FT      TRANSMEM     464   482
FT      DISULFID      152   166
FT      CARBOHYD      35    35
FT      CARBOHYD      92    92
FT      CAROXYHD     137   137
FT      CAROXYHD     165   165
FT      VARSPPLIC     17    17
FT      /length=495 000076

```

FT	CONFLICT	9	9	V -> VV (IN REF. 3).
FT	CONFLICT	164	164	MISSING (IN REF. 2).
FT	CONFLICT	230	231	KP -> NA (IN REF. 2 AND 3).
FT	CONFLICT	287	288	MISSING (IN REF. 3).
FT	CONFLICT	290	290	D -> N (IN REF. 3).
5Q	SEQUENCE	495 AA;	55860 MM;	688F37AB80448F7F CRC64;

Query Match	57.4%	Score 35;	DB 1;	Length 495;
Best Local Similarity	36.4%	Pred. No. 95;		
Matches	4;	Conservative	4;	Mismatches 3;
				Indels 0;
				Gaps 0;

```

OY      1 SLMMITOCFL 11
        ::||::|||
Db      327 TMASWKECFL 337

```

Search completed: February 10, 2004, 17:18:04
Job time : 9.34483 secs

Long).
/ETTD-YSP 000076


```

RESULT 11
ID VHY4 YEAST STANDARD; PRT: 1070 AA.
AC P38650;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 123.0 kDa protein in SP016-REC104 intergenic region.
GN YH5154W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Faveille A., Fulton L., Gattung S., Geisel C., Kirsten U.,
RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
RA Lareelle P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,
RA Nhan M., Rifkin L., Riles L., St Peter H., Trevasakis E., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VII."
RL Science 265:2077-2082 (1994).
CC -1 SIMILARITY: Contains 2 BRCT domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U10397; AAB6978.1; -.
CC PIR: S46755; S46755.
CC SGP: S0001197; BEC4.
CC GO: GO:0000335; P:negative regulation of DNA transposition; IMP.
CC InterPro: IPR001357; BRCT.
CC DR Pfam: PF00533; BRCT; 5.
CC SMART: SM00292; BRCT; 4.
CC PROSITE: PS0172; BRCT; 2.
CC KW Hypothetical protein; Repeat.
CC FT DOMAIN 117 210 BRCT 1.
CC FT DOMAIN 370 466 BRCT 2.
CC SEQUENCE 1070 AA; 123017 MW; 767931285B52580 CRC64;
SO

Query Match 59.0%; Score 36; DB 1; Length 1070;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

OX NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88160059; PubMed=2831662;
RA Bubb V., McCance D.J., Schlegel R.;
RT "DNA sequence of the HPV-16 E5 ORF and the structural conservation of
RT its encoded protein."
RT Virology 163:243-246 (1998).
RN [2]
RP SEQUENCE OF 20-83 FROM N.A.
RX MEDLINE=85246220; PubMed=2990099;
RA Seedorf K., Kramer G., Durst M., Suhai S., Rowekamp W.G.;
RT "Human papillomavirus type 16 DNA sequence."
RL Virology 145:181-185 (1985).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: K02718; AAA46938.1; ALT_SEQ.
CC PIR: A30016; W5WLS.
CC InterPro: IPR004270; Papilloma_E5.
CC DR Pfam: PF03025; Papilloma_E5; 1.
CC KW Early protein.
SO SEQUENCE 83 AA; 9401 MW; 442C0ABF0D77C0CF CRC64;

Query Match 58.2%; Score 35.5; DB 1; Length 83;
Best Local Similarity 46.7%; Pred. No. 14;
Matches 7; Conservative 3; Mismatches 0; Indels 5; Gaps 1;

ID V060 MYCPN STANDARD; PRT: 299 AA.
AC P75042;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative glycosyl transferase MG060 homolog (EC 2.-.-.-)
DE (D09 orf299).
GN MPN075 OR MP080.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfreisch R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae."
RL Nucleic Acids Res. 24:4420-4449 (1996).
CC -1 SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC STRONG, TO M.GENITALIUM MG060.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE000010; AAB95728.1; -.

```

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, Columbia;
 RX MEDLINE=21016720, PubMed=1130713;
 RA Salamouch M., Lemcke K., Rieger M., Anserge M., Unseld M.,
 RA Farman B., Valle G., Bioecker H., Perez-Alonso M., Obermaier B.,
 RA Delany M., Boutry M., Givelli L.A., Mache R., Pulgomech P.,
 RA De Simone V., Choise N., Artiguenave F., Robert C., Broctier P.,
 RA Wincker P., Catellico L., Weissenbach J., Saurin W., Quetier F.,
 RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Bens V.,
 RA Wundbach E., Drzonek H., Erle H., Jordan N., Bangert S.,
 RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nakamura G.,
 RA Vezzi A., D'Angelo M., Pallavicini A., Toppi S., Simionati B.,
 RA Conrad A., Hornischer K., Kauer R., Loehner T.-H., Nordiek G.,
 RA Reichelt J., Schaefer M., Schoen O., Barques M., Terol J., Clement J.,
 RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
 RA Cooke R., Lande M., Berger-Liauro C., Purnelle B., Masny D.,
 RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
 RA Montfort A., Agirion A., Flores M., Liguori R., Vitale D.,
 RA Manhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
 RA Mayer K.F.K., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
 RA Rooney T., Rizzo M., Walts A., Uteback T., Fujii C.Y., Shea T.P.,
 RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
 RA Pal G., Miltischer J., Sellers P., Gill J.E., Feldblyum T.V.,
 RA Preuss D., Lin X., Nieman M.C., Salzberg S.L., White O., Venter J.C.,
 RA Fraser C.W., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
 RA Sasagawa S., Kimura T., Ideawara K., Kawashima K., Kishida Y.,
 RA Kiyosawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakayama S., Nakazaki N., Nishio S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Tabata S.,
 RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
 thaliana."
 RL Nature 408:820-822(2000).
 CC -1- FUNCTION: SUBUNIT OF THE INTEGRAL MEMBRANE V0 COMPLEX OF VACUOLAR
 ATPASE. VACUOLAR ATPASE IS RESPONSIBLE FOR ACIDIFYING A VARIETY OF
 CC INTRACELLULAR COMPARTMENTS IN EUKARYOTIC CELLS, THUS PROVIDING
 CC MOST OF THE ENERGY REQUIRED FOR TRANSPORT PROCESSES IN THE
 CC VACUOLAR SYSTEM.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
 CC H(+) (out)
 CC -1- SUBUNIT: V-ATPase is a heteromultimeric enzyme composed of a
 CC peripheral catalytic V1 complex (components A to H) attached to
 CC an integral membrane V0 proton pore complex (components: a, c, c',
 CC c'', and d).
 CC -1- SIMILARITY: Belongs to the V-ATPase V0D/AC39 subunit family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: APO02057, BAB03168.1, -
 DR EMBL: APO00420, BAB03168.1, JOINED.
 DR InterPro: IPR002843, ATPase AC39sub.
 DR Pfam: PF01992, V-ATP-ATP AC39, 1.
 KM Hydrolyase; Hydrogen ion transport.
 SQ SEQUENCE 351 AA; 40787 MW; DE455CA48329544 CRC64;

RESULT 10
 ID TP53_YEAST STANDARD; PRT; 1022 AA.
 AC P38426;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Alpha, alpha-trehalose-phosphate synthase [UDP-forming] 115 kDa subunit
 DE (EC 2.4.1.15) (Trehalose-6-phosphate synthase) (UDP-glucose-
 DE glucosephosphate glucosyltransferase).
 GN TP53 OR YMR261C OR YMR156.03C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 CX NCBI_TaxID=4932;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Manning A.M., Rosenbloom C.L., Beaudet A.L.;
 RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=8288C / AB972;
 RX PubMed=9169872;
 RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
 RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
 RA Jagsels K., Iye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
 RA Rice P., Skelton J., Walsh S., Whitehead S., Barrett B.G.,
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
 XII."
 RL Nature 387:90-93(1997).
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-glucose 6-phosphate = UDP +
 CC alpha, alpha-trehalose 6-phosphate.
 CC -1- SUBUNIT: TREHALOSE SYNTHASE/PHOSPHATASE COMPLEX CONTAINS THREE
 CC OR FOUR POLYPEPTIDES OF 56 kDa (TP51), 102 kDa (TP52), 115 kDa
 CC (TP53) AND 123 kDa (TS11).
 CC -1- INDUCTION: Repressed by glucose.
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE
 CC GLYCOSYLTRANSFERASE FAMILY 20.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M88172; AAA35224.1; ALT INIT.
 DR EMBL: Z49260; CA89244.1; ALT INIT.
 DR SGD: S0004874; TP53.
 DR GO: GO:0005946; C:alpha, alpha-trehalose-phosphate synthase C...; IGI.
 DR GO: GO:0003825; F:alpha, alpha-trehalose-phosphate synthase (U...; IMP.
 DR GO: GO:0006950; P:response to stress; IMP.
 DR GO: GO:0005992; P:trehalose biosynthesis; IGI.
 DR InterPro: IPR001830, Glyco trans 20.
 DR InterPro: IPR003337, Trehalose Phase.
 DR Pfam: PF00982; Glyco transf 20, 1.
 DR Pfam: PF02358; Trehalose Phase, 1.
 KM Transferase; Glycosyltransferase; Glycolysis.
 FT DOMAIN 255 746
 FT TRANSF 154 746 P -> L (IN REF. 1).
 FT CONFLICT 733 733 N -> D (IN REF. 1).
 FT CONFLICT 802 802 S -> G (IN REF. 1).
 FT CONFLICT 870 870 D -> E (IN REF. 1).
 SQ SEQUENCE 1022 AA; 115363 MW; 49AD0A9E194E5962 CRC64;

Query Match 59.0%; Score 36; DB 1; Length 351;
 Best Local Similarity 71.4%; Pred. No. 46;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Query Match 59.0%; Score 36; DB 1; Length 1022;
 Best Local Similarity 83.3%; Pred. No. 1,3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 7
SN64 HUMAN STANDARD; PRT; 234 AA.
ID SN64 HUMAN
AC 095473;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Synaplogyrin 4.
GN SINGR4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_Taxid=9606;
[1]
RN SEQUENCE FROM N.A.
RA Kedra D., Dunamski J.P.;
RT "Cloning of a novel member of synaplogyrin gene family."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE SYNAPLOGYRIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; A2011733; CA09754.1; -
DR Genew; HENC:11502; SINGR4.
DR InterPro; IPR004904; Synaplogyrin.
DR Pfam; PF01284; MARVEL; 1.
DR Transmembrane.
KW TRANSMEM 25 45 POTENTIAL.
FT TRANSMEM 66 86 POTENTIAL.
FT TRANSMEM 104 124 POTENTIAL.
FT TRANSMEM 145 165 POTENTIAL.
SQ SEQUENCE 234 AA; 25786 MW; 801134E4D840288F CRC64;
Query Match 59.0%; Score 36; DB 1; Length 234;
Best Local Similarity 54.5%; Pred. No. 31;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
CY 1 SLMMITOCPL 11
DB 153 SILWIFQAYL 163

```

```

RESULT 8
V0D1 ARATH STANDARD; PRT; 351 AA.
ID V0D1 ARATH
AC 09LJ15;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable vacuolar ATP synthase subunit d 1 (EC 3.6.3.14) (V-ATPase d
DE subunit 1) (Vacuolar proton pump d subunit 1).
GN AT3G28710 OR MZM14.21.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_Taxid=3702;
[1]
RN SEQUENCE FROM N.A.
RA STRAIN=CV, Columbia;
RA MEDLINE=21016720; PubMed=11130713;
RA Melanoubat M., Lemcke K., Rieger M., Ansoorge W., Unseid M.,
RA Fatemoun B., Valle G., Blocker H., Perez-Alonso M., Obermaier B.,
RA Delany M., Boutry M., Grivell L.A., Mache R., Pilsdomech P.,
RA De Simone V., Choiane N., Artiguenave F., Robert C., Brottier P.,
Wincker P., Cattolico L., Weisenbach J., Saurin W., Querier F.,
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Wurmback E., Drzonek H., Erle H., Jordan N., Baergert S.,
RA Wiedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA Vezzi A., D'Angelo M., Pallavicini A., Toppi S., Simionati B.,
RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,
RA Reichelt J., Scharfe M., Schoen O., Barques M., Terol J., Clement J.,
RA Navarro P., Collado C., Perez-Perez A., Ottenwelder B., Duchemin D.,
RA Cooke R., Lautie M., Berger-Liauro C., Parnelle B., Masu D.,
RA de Haan W., Maare A.C., Alcaraz J.-P., Collet A., Casacuberta E.,
RA Montfort A., Argitrou A., Flores M., Lignori R., Vitale D.,
RA Manhaupt G., Haase D., Schoof H., Ridd S., Zaccaria P., Mewes H.-W.,
RA Mayer K.F.K., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Wals A., Utterback T., Fujii C.Y., Shea T.P.,
RA Creasy T.H., Haas B., Malt R., Wu D., Peterson J., Van Aken S.,
RA Pail G., Miltecher U., Sellers P., Gill J.E., Feldlyum T.V.,
RA Preuss D., Lin X., Niernan W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneo T., Nakamura Y., Sato S., Kato T., Aamizu E.,
RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shimo S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT thaliana."
RL Nature 408:820-822 (2000).
CC -1- FUNCTION: SUBUNIT OF THE INTEGRAL MEMBRANE V0 COMPLEX OF VACUOLAR
CC ATPASE. VACUOLAR ATPASE IS RESPONSIBLE FOR ACIDIFYING A VARIETY OF
CC INTRACELLULAR COMPARTMENTS IN EUKARYOTIC CELLS, THUS PROVIDING
CC MOST OF THE ENERGY REQUIRED FOR TRANSPORT PROCESSES IN THE
CC VACUOLAR SYSTEM.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
CC H(+) (out).
CC -1- SUBUNIT: V-ATPase is a heteromultimeric enzyme composed of a
CC peripheral catalytic V1 complex (components A to H) attached to
CC an integral membrane V0 proton pore complex (components: a, c, c',
CC c'', and d).
CC -1- SIMILARITY: Belongs to the V-ATPase V0D/AC3 subunit family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AP000420; BAB02186.1; -
DR InterPro; IPR002843; ATPsyn AC39sub.
DR Pfam; PF01992; V-ATP-synt AC39; 1.
KW Hydrolyase; Hydrogen ion transport.
SQ SEQUENCE 351 AA; 40791 MW; 5E1A67A149AC4EF4 CRC64;
Query Match 59.0%; Score 36; DB 1; Length 351;
Best Local Similarity 71.4%; Pred. No. 46;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
CY 3 LMWITOC 9
DB 328 LMWISFC 334

```

```

RESULT 9
V0D2 ARATH STANDARD; PRT; 351 AA.
ID V0D2 ARATH
AC 09LH44;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable vacuolar ATP synthase subunit d 2 (EC 3.6.3.14) (V-ATPase d
DE subunit 2) (Vacuolar proton pump d subunit 2).
GN AT3G28715 OR MZM14.22.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```


RT "A survey of the Mycoplasma genitalium genome by using random
 sequencing." 175:7918-7930(1993).
 RL J. Bacteriol. 175:7918-7930(1993).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U39704; AAC71475.1; -;
 DR EMBL: U02174; AAD12457.1; -;
 DR EMBL: U02164; AAD12446.1; -;
 DR PIR: B64228; B64228.
 DR TIGR: MG255; -;
 FT TRANSMEM 105 125 POTENTIAL.
 FT TRANSMEM 151 171 POTENTIAL.
 FT TRANSMEM 187 207 POTENTIAL.
 FT CONFLICT 180 180 N -> KR (IN REF. 2).
 SQ SEQUENCE 365 AA; 42440 MW; 6609C2C6CA2CC5B7 CRC64;
 Query Match 60.7%; Score 37; DB 1; Length 365;
 Best Local Similarity 55.6%; Pred. No. 33;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 3 LMTQCFLL 11
 DB 117 LMTQCFLL 125
 RESULT 6
 YQIA_ECOLI STANDARD; PRT; 193 AA.
 ID YQIA_ECOLI
 AC P36653
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein yqia.
 GN YQIA OR B3031 OR C3777 OR Z4388 OR ECS3919 OR SF3071.
 OS Escherichia coli.
 OS Escherichia coli O6.
 OS Escherichia coli O157:H7, and
 OS Shigella flexneri.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OC NCBI_TaxID=562, 217992, 83334, 623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=K12 / MG1655;
 RX MEDLINE=97426177 PubMed=9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12."
 RT Science 277:1453-1474(1997).
 RU [2]
 RN SEQUENCE FROM N.A.
 RP SPECIES=E.coli; STRAIN=O6:HI / CFT073 / ATCC 700928;
 RX MEDLINE=22388334; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Rosch P.,
 RA Raeko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blatner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 of uropathogenic Escherichia coli."
 RT Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 RN [3]
 RP SEQUENCE FROM N.A.

RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postel G., Hackett J., Link S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamous K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blatner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
 RT Nature 409:525-533(2001).
 RL Nature 409:525-533(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Tida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kunera S., Shiba T., Hattori W., Shinagawa H.,
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12."
 RT DNA Res. 8:11-22(2001).
 RN [5]
 RP SEQUENCE OF 1-137 FROM N.A.
 RC SPECIES=E.coli; STRAIN=K12;
 RA Imanura R., Niki H., Yamanka K., Ogura T., Fujita N., Ishihama A.,
 RA Hiraga S.;
 RT Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
 RL [6]
 RN SEQUENCE FROM N.A.
 RP SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
 RX MEDLINE=22272406; PubMed=12384590;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
 through comparison with genomes of Escherichia coli K12 and O157."
 RT Nucleic Acids Res. 30:4432-4441(2002).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U28377; AAA69189.1; -;
 DR EMBL: AE000385; AAC76067.1; -;
 DR EMBL: AE016766; AAM82221.1; -;
 DR EMBL: AE005533; AAG58170.1; -;
 DR EMBL: AP002564; BAB37342.1; -;
 DR EMBL: D16557; NOT ANNOTATED_CDS.
 DR EMBL: AB015319; AAM4549.1; -;
 DR PIR: D58723; D58723.
 DR PIR: D58723; D58723.
 DR PIR: E65090; E65090.
 DR PIR: F85963; F85963.
 DR PIR: G91118; G91118.
 DR Ecogen: EG12186; Yqia.
 DR InterPro: IPR000379; Ser esterase site.
 KM Hypothetical protein; Complete proteome.
 FT CONFLICT 135 137 ABD -> GRI (IN REF. 5).
 SQ SEQUENCE 193 AA; 21641 MW; 3077DE58EC058A3 CRC64;
 Query Match 59.0%; Score 36; DB 1; Length 193;
 Best Local Similarity 57.1%; Pred. No. 26;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 5 WITQCFLL 11
 DB 77 WITQCFLL 83

FT TRANSMEM 2496 2516 POTENTIAL.
 FT DOMAIN 2517 2670 CYTOPLASMIC (POTENTIAL).
 FT MOD RES 2582 2582 PHOSPHORYLATION (POTENTIAL).
 SQ SEQUENCE 2670 AA; 304284 MW; 13C787E4C2886E45 CRC64;

Query/Match 62.3%; Score 38; DB 1; Length 2670;
 Best Local Similarity 70.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SLIMWITQCF 10
 DB 2239 SLIFWILICF 2248

RESULT 4
 IP3T HUMAN STANDARD; PRT; 2671 AA.
 AC 014573; 014649;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Inositol 1,4,5-trisphosphate receptor type 3 (Type 3 inositol 1,4,5-
 DE trisphosphate receptor) (Type 3 InsP3 receptor) (IP3 receptor isoform
 DE 3) (InsP3R3).
 GN ITPR3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RX MEDLINE=94363219; PubMed=8081734;
 RA Yamamoto-Hino M., Sugiyama T., Hikita K., Mattei M.-G., Hasegawa K.,
 RA Sekine S., Sakurada K., Miyawaki A., Furuchi T., Hasegawa M.,
 RA Mochizuki K., and characterization of human type 2 and type 3 inositol
 RT "Cloning and characterization of human type 2 and type 3 inositol
 RT 1,4,5-trisphosphate receptors."
 RL Recept. Channels 2:9-22(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94117432; PubMed=8288584;
 RA Maranto A.R.;
 RT Primary structure, ligand binding, and localization of the human type
 RT 3 inositol 1,4,5-trisphosphate receptor expressed in intestinal
 RT epithelium."
 RL J. Biol. Chem. 269:1222-1230(1994).
 CC -1- FUNCTION: RECEPTOR FOR INOSITOL 1,4,5-TRISPHOSPHATE. A SECOND
 CC -1- MESSENGER THAT MEDIATES THE RELEASE OF INTRACELLULAR CALCIUM.
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY). INTERACTS WITH TRP1,
 CC -1- TRPC3 AND TRPC4.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN INTESTINAL CRYPT AND VILLUS
 CC -1- BETHLEHEM CELLS.
 CC -1- DOMAIN: THE RECEPTOR CONTAINS A CALCIUM CHANNEL IN ITS C-TERMINAL
 CC -1- EXTREMITY. ITS LARGE N-TERMINAL CYTOPLASMIC REGION HAS THE LIGAND-
 CC -1- BINDING SITE IN THE N-TERMINUS AND MODULATORY SITES IN THE MIDDLE
 CC -1- PORTION IMMEDIATELY UPSTREAM OF THE CHANNEL REGION.
 CC -1- PFM: PHOSPHORYLATED ON TYROSINE RESIDUES (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INSP3 RECEPTOR FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D26351; BA05385.1; -
 CC EMBL; U01062; AAC50064.1; -
 CC PIR; A49873; A49873.
 CC Genew; HGNC:6182; ITPR3.

DR MIM; 147267; -
 DR InterPro; IPR000699; Ca-rel_channel.
 DR InterPro; IPR001682; Ca/Na_pore.
 DR InterPro; IPR005821; Ion_trans.
 DR InterPro; IPR003608; MIR.
 DR Pfam; PF00520; Ion_trans; 1.
 DR Pfam; PF02815; MIR; 4.
 DR Pfam; PF01365; MYR_ITPR; 2.
 DR SMART; SM00472; MIR; 4.
 DR Receptor; Transmembrane; Phosphorylation; Endoplasmic reticulum;
 KW Receptor; Transmembrane; Ion transport; Calcium channel.
 KM Ionic channel; Ion transport; Calcium channel.
 FT DOMAIN 1 2202
 FT TRANSMEM 2203 2223 POTENTIAL.
 FT DOMAIN 2224 2235 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 2236 2256 POTENTIAL.
 FT DOMAIN 2257 2264 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 2265 2285 POTENTIAL.
 FT TRANSMEM 2286 2325 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 2326 2346 POTENTIAL.
 FT TRANSMEM 2347 2368 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 2369 2389 POTENTIAL.
 FT DOMAIN 2390 2496 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 2497 2517 POTENTIAL.
 FT DOMAIN 2518 2671 CYTOPLASMIC (POTENTIAL).
 FT MOD RES 2583 2583 PHOSPHORYLATION (POTENTIAL).
 FT TRANSMEM 524 524 A -> V (IN REF. 2).
 FT CONFLICT 562 562 H -> Y (IN REF. 2).
 FT CONFLICT 989 989 H -> Y (IN REF. 2).
 FT CONFLICT 1143 1143 A -> T (IN REF. 2).
 FT CONFLICT 1391 1391 L -> V (IN REF. 2).
 FT CONFLICT 1496 1497 L -> V (IN REF. 2).
 FT CONFLICT 1674 1674 L -> V (IN REF. 2).
 FT CONFLICT 2436 2436 V -> L (IN REF. 2).
 SQ SEQUENCE 2671 AA; 304036 MW; 69F618CF27B681D7 CRC64;

Query/Match 62.3%; Score 38; DB 1; Length 2671;
 Best Local Similarity 70.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SLIMWITQCF 10
 DB 2240 SLIFWILICF 2249

RESULT 5
 Y55 MYCGE STANDARD; PRT; 365 AA.
 ID Y55 MYCGE 049289; 049297;
 AC P47457; 049289; 049297;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein MG255.
 GN MG255.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2097;
 RN [1]
 RP SEQUENCE FROM N.A. / G-37;
 RC STRAIN=ATCC 33530 / G-37;
 RX MEDLINE=96026346; PubMed=756993;
 RA Frazer C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
 RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
 RA Tomby J.F., Dougherty B.A., Bott K.F., Hu P.-C., Ucker T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RT "The minimal gene complement of Mycoplasma genitalium."
 RL Science 270:397-403(1995).
 RN [2]
 RP SEQUENCE OF 134-365 FROM N.A.
 RC STRAIN=ATCC 33530 / G-37;
 RX MEDLINE=94075230; PubMed=8253680;
 RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;

```

FT DOMAIN 5 82 GLY-RICH.
FT TRANSMEM 156 172 POTENTIAL.
SQ SEQUENCE 180 AA; 17992 MW; B122C5C2C8BE1569 CRC64;

Query Match 100.0%; Score 61; DB 1; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLMMWITOCFL 11
DB 157 SLMMWITOCFL 167

RESULT 2
ATP6_BUCAL STANDARD; PRT; 274 AA.
AC P57118;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP synthase A chain (EC 3.6.3.14) (Protein 6).
GN ATP6 OR BU002.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OC symbiotic bacterium).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT Buchnera sp. APS.;
RL Nature 407:81-86(2000).
CC -1- FUNCTION: KEY COMPONENT OF THE PROTON CHANNEL; IT MAY PLAY A
CC DIRECT ROLE IN THE TRANSLLOCATION OF PROTONS ACROSS THE MEMBRANE
CC (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
CC H(+) (out)
CC -1- SUBUNIT: P-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC HAS THREE MAIN SUBUNITS: A, B AND C (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1- SIMILARITY: BELONGS TO THE ATPASE A CHAIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP001118; BAB12730.1; -
CC HSSP; P00855; 1C17
CC InterPro; IPR000568; ATPaseN Asub.
CC Pfam; PF00119; ATP-asynt A; 1
CC PRINTS; PR00123; ATPASEA.
CC TIGR; TIGR01131; ATP_synt_6_of_1.
DR PROSITE; PS00449; ATPase_A_1.
KM Hydrogen ion transport; CF(0); Transmembrane; Complete proteome.
FT TRANSMEM 44 64 POTENTIAL.
FT TRANSMEM 110 130 POTENTIAL.
FT TRANSMEM 156 176 POTENTIAL.
FT TRANSMEM 179 199 POTENTIAL.
FT TRANSMEM 212 232 POTENTIAL.
FT TRANSMEM 243 263 POTENTIAL.
SQ SEQUENCE 274 AA; 31363 MW; 4B2321336FE78A4 CRC64;

Query Match 70.5%; Score 43; DB 1; Length 274;
Best Local Similarity 80.0%; Pred. No. 2.4;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 LMMWITOCFL 11
DB 229 LMMWITOCFL 238

RESULT 3
IP3T_RAT STANDARD; PRT; 2670 AA.
AC Q63259;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Inositol 1,4,5-trisphosphate receptor type 3 (Type 3 inositol 1,4,5-
DE trisphosphate receptor) (Type 3 InsP3 receptor) (IP3 receptor isoform
DE 3) (InsP3R3) (IP3R-3).
GN IP3R3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9326594; PubMed=8388391;
RA Blondel O., Takeda J., Janssen H., Seino S., Bell G.I.;
RT "Sequence and functional characterization of a third inositol
RT trisphosphate receptor subtype, IP3R-3, expressed in pancreatic
RT islets, kidney, gastrointestinal tract, and other tissues.";
RL J. Biol. Chem. 268:11356-11363(1993).
CC -1- FUNCTION: RECEPTOR FOR INOSITOL 1,4,5-TRISPHOSPHATE, A SECOND
CC MESSENGER THAT MEDIATES THE RELEASE OF INTRACELLULAR CALCIUM.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum (By similarity).
CC -1- DOMAIN: THE RECEPTOR CONTAINS A CALCIUM CHANNEL IN ITS C-TERMINAL
CC EXTREMITY. ITS LARGE N-TERMINAL CYTOPLASMIC REGION HAS THE LIGAND-
CC BINDING SITE IN THE N-TERMINUS AND MODULATORY SITES IN THE MIDDLE
CC PORTION IMMEDIATELY UPSTREAM OF THE CHANNEL REGION.
CC -1- PTM: PHOSPHORYLATED ON TYROSINE RESIDUES (BY SIMILARITY).
CC
CC -1- SIMILARITY: BELONGS TO THE INSP3 RECEPTOR FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L06096; AAA4146.1; -
CC PIR; A46719; A46719.
DR InterPro; IPR000699; Ca-rel channel.
DR InterPro; IPR001682; Ca/Na pore.
DR InterPro; IPR005821; Ion_pore.
DR InterPro; IPR003608; MIR.
DR Pfam; PF00520; Ion_trans; 1.
DR Pfam; PF02815; MIR; 4.
DR Pfam; PF01365; RYDR_IP3R; 2.
DR SMART; SMO0472; MIR; 4.
KW Receptor; Transmembrane; Phosphorylation; Endoplasmic reticulum;
KW Ionic channel; Ion transport; Calcium channel.
FT TRANSMEM 1 2201 POTENTIAL.
FT TRANSMEM 2202 2222 POTENTIAL.
FT TRANSMEM 2223 2233 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 2234 2254 POTENTIAL.
FT TRANSMEM 2255 2263 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 2264 2284 POTENTIAL.
FT TRANSMEM 2285 2324 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 2325 2345 POTENTIAL.
FT TRANSMEM 2346 2367 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 2368 2388 POTENTIAL.
FT DOMAIN 2389 2495 EXTRACELLULAR (POTENTIAL).

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 17:13:40 ; Search time 8.34483 Seconds
(without alignments)
61.990 Million cell updates/sec

Title: US-10-023-182-4
Perfect score: 61
Sequence: 1 SLWMITQCP 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	100.0	180	CTG1_HUMAN	P78358 homo sapien
2	43	70.5	274	ATP6_BUCAL	P57118 buchnera ap
3	38	62.3	2670	IP3T_RAT	Q63269 rattus norv
4	38	62.3	2671	IP3T_HUMAN	Q14573 homo sapien
5	37	60.7	365	Y255_MYCOE	P47497 mycoplasma
6	36	59.0	193	VO1A_ECOLI	P36653 escherichia
7	36	59.0	234	SNG4_HUMAN	Q95473 homo sapien
8	36	59.0	351	VOD1_ARATH	Q91145 arabidopsis
9	36	59.0	351	VOD2_ARATH	Q91146 arabidopsis
10	36	59.0	1022	TPS3_YEAST	P38426 saccharomyc
11	36	59.0	1070	YHV4_YEAST	P38850 saccharomyc
12	35.5	58.2	83	VE5_HPV16	P06972 human papil
13	35	57.4	299	POG2_MYCPN	P55042 mycoplasma
14	35	57.4	341	PCG2_CAYPO	Q60513 cavia porce
15	35	57.4	495	ACHP_RAT	P23322 rattus norv
16	35	57.4	693	RAP1_SCHPO	Q96177 schistosom
17	35	57.4	806	RIR1_HSVJ	P06641 human herpe
18	35	57.4	1021	ALAI_CHICK	P09572 gallus gall
19	35	57.4	1458	CUT4_SCHPO	Q91974 schistosom
20	35	57.4	1950	UBR1_YEAST	P34451 unknknem v
21	35	57.4	2103	RRP1_YUK	Q03880 saccharomyc
22	34	55.7	122	YML0_YEAST	P06324 mus musculu
23	34	55.7	132	TVG3_MOUSE	P04435 homo sapien
24	34	55.7	133	TVB2_MOUSE	P01740 mus musculu
25	34	55.7	135	TVCI_MOUSE	P11974 fowlpox vir
26	34	55.7	219	DKK1_FOXPV	Q91972 homo sapien
27	34	55.7	309	TRR8_HUMAN	Q91972 homo sapien
28	34	55.7	309	TRR9_RAT	Q91972 homo sapien
29	34	55.7	334	Y278_METH	Q06472 methanobact
30	34	55.7	345	CAAR_CAYPO	Q07012 cavia porce
31	34	55.7	353	YRRI_BACSU	Q04472 bacillus su
32	34	55.7	385	YERY_ECOLI	P33161 escherichia
33	34	55.7	504	PSD3_CABEL	Q04908 caenorhabdi

ALIGNMENTS

RESULT 1	CTG1_HUMAN	STANDARD	PRT	180 AA.
ID	CTG1_HUMAN			
AC	P78358			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Cancer/testis antigen 1 (Autoimmunogenic cancer/testis antigen NY-ESO-1).			
GN	CTAG1 OR CTAG.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97203161; PubMed=9050879;			
RA	Chen Y.-T., Scanlan M.U., Sahin U., Tureci O., Gure A.O., Tsang S.,			
RA	Williamson B., Stockert E., Pfundschuh W., Old U.U.;			
RT	"A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening."			
RL	Proc. Natl. Acad. Sci. U.S.A. 94:1914-1918 (1997).			
RL	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Melanoma;			
RX	MEDLINE=98289662; PubMed=9626360;			
RA	Lette B., Lucas S., Michaux L., de Smet C., Godelaine D., Serrano A.,			
RA	de Plaen E., Boon T.;			
RT	"LAG-1, a new gene with tumor specificity."			
RL	Int. J. Cancer 76:903-908 (1998).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98430682; PubMed=9759882;			
RA	Wang R.-F., Johnston S.L., Zeng G., Topalian S.L.,			
RA	Schwartzentruber D.U., Rosenberg S.A.;			
RT	"A breast and melanoma-shared tumor antigen: T cell responses to			
RT	antigenic peptides translated from different open reading frames."			
RL	J. Immunol. 161:3596-3606 (1998).			
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN TESTIS AND OVARY AND IN A WIDE			
CC	VARIETY OF CANCERS. DETECTED IN UTERINE MYOMETRIUM.			
CC	-1- SIMILARITY: BELONGS TO THE CTAG FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
DR	EMBL: U87459; AAB49693.1; -			
DR	EMBL: AJ003149; CA05908.1; -			
DR	EMBL: AF038567; AAD05202.1; -			
DR	GeneW: HGNC:2491; CTAG1.			
DR	MIM: 300156; -			
DR	Transmembrane; Antigen.			

34	34	55.7	513	1	MAMB_BACSM	P16639 bacillus sp
35	34	55.7	669	1	DAFI_CABEL	P20732 caenorhabdi
36	34	55.7	1502	1	MRP6_RAT	O88269 rattus norv
37	34	55.7	1503	1	MRP6_HUMAN	O95255 homo sapien
38	34	55.7	3110	1	HD_RAT	P51111 rattus norv
39	34	55.7	3119	1	HD_MOUSE	P42859 mus musculu
40	34	55.9	1545	1	MRP2_HUMAN	Q92887 homo sapien
41	33	54.1	114	1	APR2_HUMAN	Q95511 homo sapien
42	33	54.1	163	1	PHZA_PSEPL	Q51787 pseudomonas
43	33	54.1	213	1	CYB6_HEIGE	Q91598 heliobacter
44	33	54.1	215	1	CIB3_HUMAN	Q91972 homo sapien
45	33	54.1	237	1	R33A_HUMAN	Q14068 homo sapien

., Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodex, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L. Nature 390, 364-370, 1997
 A:Authors: Uteerback, T.; Cotton, M.D.; Spriggs, T.; Artlach, P.; Kaine, B.P.; Sykes, S. Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: E69322
 A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-581 <KLE>
 A:Cross-references: GB:AE001064; GB:AE000782; NID:G2689387; PIDN:AA90655.1; PID:G265003

Query Match 59.8%; Score 36.5; DB 2; Length 581;
 Best Local Similarity 57.1%; Pred. No. 1e+02;
 Matches 8; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

OY 1 SLMM--ITQCF 11
 |||||
 Db 479 SLMLMSDITVCYL 492

RESULT 13
 D58723
 hypothetical protein (orf4, cpdA 5' region) - Escherichia coli
 C:Species: Escherichia coli
 C:Date: 18-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 20-Feb-1998
 C:Accession: D58723
 R:Ramura, R.; Yamanaka, K.; Ooura, T.; Hiraga, S.; Fujita, N.; Ishihama, A.; Niki, H. J. Biol. Chem. 271, 25423-25429, 1996
 A:Title: Identification of the cpdA gene encoding cyclic 3',5'-adenosine monophosphate P
 A:Reference number: A58723; MUID:96411758; PMID:8810311
 A:Accession: D58723
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-137 <IMA>
 A:Cross-references: GB:D16557; NID:G453393
 A:Experimental source: strain K-12
 C:Genetics:
 A:Map position: 68.4 min

Query Match 59.0%; Score 36; DB 2; Length 137;
 Best Local Similarity 57.1%; Pred. No. 35;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 5 WITQCF 11
 |||||
 Db 77 WLSQCFM 83

RESULT 14
 F85963
 hypothetical protein yqIA [imported] - Escherichia coli (strain O157:H7, substrain EDL93
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: F85963
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grobeck, E.T.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamovotis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85460; MUID:21074935; PMID:11206551
 A:Accession: F85963
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-193 <STO>
 A:Cross-references: GB:AE005174; NID:G12517609; PIDN:AA958170.1; GSPDB:GN00145; TMGP:243
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: yqIA

Query Match 59.0%; Score 36; DB 2; Length 193;
 Best Local Similarity 57.1%; Pred. No. 47;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 5 WITQCF 11
 |||||
 Db 77 WLSQCFM 83

RESULT 15
 E65090
 hypothetical 15.2 kD protein in loc 3' region - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
 C:Accession: E65090
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.
 A.; Rose, D.J.; Mau, B.; Sha, Y. Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: E65090
 A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-193 <BLAT>
 A:Cross-references: GB:AE000385; GB:U00096; NID:G1789405; PIDN:AACT6067.1; PID:G1789405
 A:Experimental source: strain K-12, substrain MG1655
 C:Genetics:
 A:Gene: yqIA

Query Match 59.0%; Score 36; DB 2; Length 193;
 Best Local Similarity 57.1%; Pred. No. 47;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 5 WITQCF 11
 |||||
 Db 77 WLSQCFM 83

Search completed: February 10, 2004, 17:29:45
 Job time: 15.2759 secs

```

A/Accession: F66410
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-996 <STO>
A/Cross-references: GB:AE005172; NID:96560764; PIDN:AAAF6764.1; GSPDB:GN00141
C/Genetics:
A/Gene: F3M18.12
A/Map position: 1
C/Superfamily: protein kinase Xaz1, leucine-rich alpha-2-glycoprotein repeat homology; P

Query Match      62.3%; Score 38; DB 2; Length 996;
Best Local Similarity 55.6%; Pred. No. 89;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Cy 3 LHMWITQCF 11
    |||:|:|
Db 209 VMMWITCFL 217

RESULT 8
A46719
C/Species: Rattus norvegicus (Norway rat)
C/Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
A/Accession: A46719
R/Blondel, O.; Takeda, J.; Janssen, H.; Setno, S.; Bell, G.I.
J. Biol. Chem. 269, 11356-11363, 1993
A/Title: Sequence and functional characterization of a third inositol trisphosphate rece
A/Reference number: A46719; MUID:93266594; PMID:8388391
A/Accession: A46719
A/Status: preliminary
A/Molecule type: nucleic acid
A/Residues: 1-2670 <ELO>
A/Cross-references: GB:L06096; NID:9310170; PIDN:AAA4446.1; PID:9310171
A/Experimental source: insulinoma cell line RINm5F
A/Note: sequence extracted from NCBI backbone (NCBIN:132840, NCBI:P.132841)
C/Superfamily: inositol-trisphosphate receptor

Query Match      62.3%; Score 38; DB 2; Length 2670;
Best Local Similarity 70.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Cy 1 SLWMWITQCF 10
    |||:|:|
Db 2239 SLFWITLICF 2248

RESULT 9
A49873
C/Species: Homo sapiens (man)
C/Date: 30-Jun-1995 #sequence_revision 30-Jun-1995 #text_change 20-Aug-1999
A/Accession: A49873
R/Marranco, A.R.
J. Biol. Chem. 269, 1222-1230, 1994
A/Title: Primary structure, ligand binding, and localization of the human type 3 inositol
A/Reference number: A49873; MUID:94117432; PMID:8286584
A/Accession: A49873
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-2671 <MAR>
A/Cross-references: GB:U01062; NID:9453367; PIDN:AAAC50064.1; PID:9393036
A/Gene: GDB:ITPR3
A/Cross-references: GDB:127551; OMIM:147267
A/Map position: 6pter-6p21
C/Superfamily: inositol-trisphosphate receptor

Query Match      62.3%; Score 38; DB 2; Length 2671;
Best Local Similarity 70.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Cy 1 SLWMWITQCF 10

```

```

Db 2240 SLFWITLICF 2249

RESULT 10
I38320
C/Species: Homo sapiens (man)
C/Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 21-Jan-2000
A/Accession: I38320
R/Slightom, J.L.; Siemieniuk, D.R.; Siew, L.C.; Koop, B.F.; Hood, L.
Genomics 20, 149-168, 1994
A/Title: Nucleotide sequence analysis of 77.7 kb of the human V beta T-cell receptor ge
A/Reference number: A54302; MUID:94292194; PMID:8020962
A/Accession: I38320
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-113 <RBS>
A/Cross-references: EMBL:U03115; NID:9467918; PIDN:AAAL7719.1; PID:9467929
C/Genetics:
A/Gene: TCRBV26S1
A/Introns: 17/1
C/Superfamily: immunoglobulin V region; immunoglobulin homology
C/Keywords: T-cell receptor
P:35-113/Domain: immunoglobulin homology <IMM>

Query Match      60.7%; Score 37; DB 2; Length 113;
Best Local Similarity 60.0%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Cy 2 LHMWITQCF 11
    |||:|:|
Db 5 LHMWITCFL 14

RESULT 11
B64228
C/Species: Mycoplasma genitalium
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 07-Dec-1999
A/Accession: B64228
R/Praeger, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.
M.; Fuhmann, J.; Nguyen, D.; Uterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.
C.A.; Venter, J.C.
Science 270, 397-403, 1995
A/Title: The minimal gene complement of Mycoplasma genitalium.
A/Reference number: A64200; MUID:96026346; PMID:7565993
A/Accession: B64228
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-365 <TIGR>
A/Cross-references: GB:U39703; GB:L43967; NID:91045933; PID:91045946; TIGR:MG255
A/Experimental source: strain G-37
C/Genetics:
A/Genetic code: SG3
C/Superfamily: Mycoplasma hypothetical protein MG255

Query Match      60.7%; Score 37; DB 1; Length 365;
Best Local Similarity 55.6%; Pred. No. 55;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Cy 3 LHMWITQCF 11
    |||:|:|
Db 117 LHMWITCFL 125

RESULT 12
B69322
C/Species: Archaeoglobus fulgidus
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
A/Accession: B69322
R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

```

A:Map position: circular chromosome
C:Superfamily: uncharacterized conserved protein

Query Match 67.2% Score 41; DB 2; Length 196;
Best Local Similarity 70.0% Pred. No. 6.5;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLMMWITQCF 10
|||:|
23 SLMMWITPAF 32

RESULT 3
G97652
hypothetical protein AGR_C_4424 [imported] - Agrobacterium tumefaciens (strain C58, Cere

C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Dec-2002
C:Accession: G97652
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A:Reference number: A97359; MUID:2160851; PMID:11743194

A:Accession: G97652

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-198 <KTR>

A:Cross-references: GB:AE007869; PIDN:AAK8176.1; PID:G15157620; GSPDB:GN00169

A:Gene: AGR_C_4424

A:Map position: circular chromosome

C:Superfamily: uncharacterized conserved protein

Query Match 67.2% Score 41; DB 2; Length 198;
Best Local Similarity 70.0% Pred. No. 6.5;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SLMMWITQCF 10
|||:|
25 SLMMWITPAF 34

RESULT 4
S38469
hypothetical protein D65 - Desulfurococcus mobilis (fragment)

C:Species: Desulfurococcus mobilis

C:Date: 13-Jan-1995 #sequence_revision 25-Oct-1996 #text_change 21-Jul-2000

C:Accession: S54739; S38469

R:Deccarelli, B.; Bochetto, M.; Creti, R.; Sanangelantoni, A.M.; Tiboni, O.; Cammarano, Mol. Gen. Genet. 246, 687-696, 1995

A:Title: Chromosomal organization and nucleotide sequence of the genes for elongation fa

obilis.

A:Reference number: S54733; MUID:95206243; PMID:7898436

A:Accession: S54739

A:Status: preliminary; translation not shown

A:Molecule type: DNA

A:Residues: 1-65 <CE2>

A:Cross-references: EMBL:X73582; NID:9410438; PIDN:CAA51989.1; PID:9410444

Query Match 62.3% Score 38; DB 2; Length 65;
Best Local Similarity 60.0% Pred. No. 8.1;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LLMWITQCF 11
|||:|
6 LLMWITQCF 15

RESULT 5
C82178
conserved hypothetical protein VC1609 [imported] - Vibrio cholerae (strain N16961 serogr

C:Accession: C82178

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwim, M.L.; Dodson, R.J.

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoti, I.; Sellers,

J. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: C82178

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-408 <HEI>

A:Cross-references: GB:AE004239; GB:AE003852; NID:9656130; PIDN:AAF94763.1; GSPDB:GN00

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC1609

A:Map position: 1

Query Match 62.3% Score 38; DB 2; Length 408;
Best Local Similarity 62.5% Pred. No. 41;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 LLMWITQCF 10
|||:|
345 LLMWITQCF 352

RESULT 6
T25899
hypothetical protein T20F7.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T25899

R:Miller, N.; Gattung, S.

submitted to the EMBL Data Library, April 1997

A:Description: The sequence of C. elegans cosmid T20F7.

A:Reference number: Z20107

A:Accession: T25899

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-478 <MLL>

A:Cross-references: EMBL:U97550; PIDN:AA852856.1; GSPDB:GN00028; CESP:T20F7.6

A:Experimental source: strain Bristol N2; clone T20F7

C:Genetics:

A:Gene: CESP:T20F7.6

A:Map position: X

A:Intron: 33/3; 112/2; 144/1; 205/3; 263/2; 300/3; 380/3; 402/2; 451/1

Query Match 62.3% Score 38; DB 2; Length 478;
Best Local Similarity 62.5% Pred. No. 47;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 LLMWITQCF 10
|||:|
43 LLMWITQCF 50

RESULT 7
F86410
protein F3M8.12 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002

C:Accession: F86410

R:Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Hulzar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marziani

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 17:17:35, Search time 13.2759 Seconds
(without alignments)
79.683 Million cell updates/sec

Title: US-10-023-182-4

Perfect score: 61

Sequence: 1 SLWMITQCF 11

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	70.5	274	2	B84930 H+-transporting tw
2	41	67.2	196	2	AE2876 conserved hypothet
3	41	67.2	198	2	G97652 hypothetical prote
4	38	62.3	65	2	S38469 hypothetical prote
5	38	62.3	408	2	C82178 conserved hypothet
6	38	62.3	478	2	T25899 hypothetical prote
7	38	62.3	996	2	F86410 protein F86410
8	38	62.3	2670	2	A46719 inositol 1,4,5-tri
9	38	62.3	2671	2	A49873 inositol 1,4,5-tri
10	37	60.7	113	2	T-cell receptor be
11	37	60.7	113	2	T-cell receptor be
12	36.5	59.8	581	2	B63222 hypothetical prote
13	36	59.0	137	2	D58723 hypothetical prote
14	36	59.0	193	2	F85963 hypothetical prote
15	36	59.0	193	2	B65090 hypothetical prote
16	36	59.0	193	2	G91118 hypothetical prote
17	36	59.0	193	2	AF0889 conserved hypothet
18	36	59.0	288	2	G90368 conserved hypothet
19	36	59.0	351	2	T03153 hypothetical prote
20	36	59.0	383	2	J03472 brain and reproduc
21	36	59.0	861	2	S77086 hypothetical prote
22	36	59.0	1054	2	S84473 TP63 protein - Yea
23	36	59.0	1070	2	S46755 hypothetical prote
24	35.5	58.2	83	1	WSW185 E5 protein - human
25	35	57.4	142	2	G96527 succinate dehydrog
26	35	57.4	242	2	G96527 protein F27J15.18
27	35	57.4	247	2	C97318 uncharacterized
28	35	57.4	270	2	A34365 Fe-gamma receptor
29	35	57.4	270	2	C97318
30	35	57.4	270	2	C97318

30	35	57.4	316	2	S50336 NADH2 dehydrogenas
31	35	57.4	380	2	A83458 hypothetical prote
32	35	57.4	483	2	T03909 hypothetical prote
33	35	57.4	492	2	S74338 glycylate oxidase
34	35	57.4	495	2	B35721 nicotinic acetylch
35	35	57.4	517	2	A30992 probable nicotinic
36	35	57.4	577	2	T09024 proline-rich prote
37	35	57.4	612	2	T39684 hypothetical prote
38	35	57.4	806	2	T41930 ribonucleotide red
39	35	57.4	974	2	A86263 F13K23.18 protein
40	35	57.4	1021	2	A28199 Na+/K+-exchanging
41	35	57.4	1458	2	T51995 hypothetical prote
42	35	57.4	1458	2	T39266 probable cell cycl
43	35	57.4	1950	2	S12332 ubiquitin-protein
44	35	57.4	2103	1	J01621 genome polypeptid
45	34	55.7	80	2	AB0913 conserved hypothet

ALIGNMENTS

RESULT 1
B84930 H+-transporting two-sector ATPase (EC 3.6.3.14) A chain [imported] - Buchnera sp. (stra)
C/Species: Buchnera sp.
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 03-Jun-2002
C/Accession: B84930
R/Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A/Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. ?
A/Reference number: A84930; MIMD:2045173; PMID:10993077
A/Accession: B84930
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-274 <STO>
A/Cross-references: GB:AP000398; GSPDB:GN00144
A/Experimental source: strain Aps
C/Genetics:
A/Genes: atpB; BU002
C/Superfamily: H+-transporting ATP synthase, protein 6
C/Keywords: hydrolase

Query Match 70.5%; Score 43; DB 2; Length 274;
Best Local Similarity 80.0%; Prod. No. 3.9;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 SLWMITQCF 11
DB 229 LIPMWITQCF 238

RESULT 2
AE2876 conserved hypothetical protein Atu2439 [imported] - Agrobacterium tumefaciens (strain C)
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Dec-2002
C/Accession: AE2876
R/Wood, D.W.; Seubal, U.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woc, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; MIMD:21608550; PMID:11743193
A/Accession: AE2876
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-196 <KUR>
A/Cross-references: GB:AE008688; PIND:AA143427.1; PID:g17740928; GSPDB:GN00186
A/Experimental source: strain C58 (Dupont)
C/Genetics:


```

/ LENGTH: 210
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-117-937-76

Alignment Scores:
Pred. No.: 7.52e-33 Length: 210
Score: 597.00 Matches: 121
Percent Similarity: 68.12% Conservative: 20
Best Local Similarity: 58.45% Mismatches: 38
Query Match: 42.55% Indels: 28
DB: 12 Gaps: 4

US-10-023-182-1 (1-752) x US-10-117-937-76 (1-210)

QY 54 ATGCAGGCGGAGGCGGCGGCAAGGCGGCTTCAGCGGCGATGCTGATGCGCCAGAGGC 113
Db 1 MetGlnAlaGluGlyArgGlyThrGlyGlySerThrGlyAspAlaAspGlyProGlyGly 20

QY 114 CTTGCGATTCCTGATGAGCCAGGCGGCAATGCTGCGCGCGCCAGAGAGGCGGTCGCG 173
Db 21 ProGlyIleProAspGlyProGlyGlyAsnAlaGlyGlyProGlyGlyAlaGlyAlaThr 40

QY 174 GCGCGCAGAGTCCCGCGGCGCGAGGCGGCGGCGCTCGGCGCGCGGCGGAGAGCGCC 233
Db 41 GlyGlyArgGlyProArgGlyAlaGlyAlaAlaArgAlaSerGlyProArgGlyGlyAla 60

QY 234 CCGCGGCGTCCCGATGCGCGCGCGCGCTTCAGGCGTGAATGCTGCTGCAATGCGGCGCC 293
Db 61 ProAlaGlyProAlaGlyAlaAlaAlaSerAlaGlnAspGlyArgGlyProGlyGlyAla 80

QY 294 AGCGGCGCGGAGAGCGGCGCTGCTGAGTTCTACCTCGCCATGCTTTTGCGGCAACCGATG 353
Db 81 ArgAlaProAspSerArgIleuGluLeuHisIleThrMetProPheSerSerProMet 100

QY 354 GAAGCAGAGCTGCGCGCGAGAGCTGCGCGCGAGATGCCCGCTTCGCGCGAGG 413
Db 101 GluAlaGluLeuValArgArgIleuSerArgAlaAlaProLeuProArgProGly 120

QY 414 GTCCTTGAAGAGAGTCACTGTCGCGGCAACATGCACTGCACTGCTGCTGCA 473
Db 121 AlaValLeuLysAspSerThrValSerGlyAsnLeuPheMetSerValIleTrpAspGln 140

QY 474 GACCAACCGCACTGCACTGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTG 533
Db 141 AspArgGluGly-----AlaGlyArgMetArgValAlaGly 152

QY 534 TGGATCAGCAGAGTCTTCTGCGCGGTTTGGCTCAGCGCTCCTCAGGAGAGG--- 590
Db 153 TTP-----GlyLeuGlySerAlaSerProGluGlyGlnLysAla 165

QY 591 -----CGCTAAGCCAGCAGCTGCGCGCGCTTCCTCTAG 620
Db 166 ArgAspLeuArgThrProLysHisLysValSerGluGlnArgProGlyThrProLys--- 184

QY 621 GTCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 641
Db 185 ---ProProProProGluGly 190

RESULT 15
US-10-342-331-3
/ Sequence 3: Application US//10342331
/ Publication No. US20030229205A1
/ GENERAL INFORMATION:
/ APPLICANT: VAN HEERDE, GEORGE V.
/ APPLICANT: VAN RIJN, ALEXIS C.
/ APPLICANT: BOWMSTRA, JAN B.
/ APPLICANT: DE WOLF, FREDERIK A.
/ APPLICANT: MOORBOEK, ANDREAS
/ APPLICANT: WERTEN, MARC W.T.
/ APPLICANT: MIND, RICHELIE D.
/ APPLICANT: VAN DEN BOSCH, TANJA J.
/ TITLE OF INVENTION: SILVER HALIDE EMULSIONS WITH RECOMBINANT COLLAGEN

```

```

/ TITLE OF INVENTION: SUITABLE FOR PHOTOGRAPHIC APPLICATION AND ALSO THE
/ TITLE OF INVENTION: PREPARATION THEREOF
/ FILE REFERENCE: 2728-2
/ CURRENT APPLICATION NUMBER: US/10/342,331
/ PRIOR FILING DATE: 2003-01-15
/ PRIOR APPLICATION NUMBER: US/09/219,849
/ PRIOR FILING DATE: 1998-12-23
/ NUMBER OF SEQ ID NOS: 50
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 3
/ LENGTH: 504
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Illustrative
US-10-342-331-3

Alignment Scores:
Pred. No.: 6.05e-06 Length: 504
Score: 202.50 Matches: 66
Percent Similarity: 43.64% Conservative: 6
Best Local Similarity: 40.00% Mismatches: 62
Query Match: 14.48% Indels: 31
DB: 12 Gaps: 10

US-10-023-182-1 (1-752) x US-10-342-331-3 (1-504)

QY 426 CTTGAGAGACCCCTTGCGACGCGGAGCGGCGGCGATCTGGG---CCAGGCTCTGCG 370
Db 212 ProAlaGlyProProGlyAlaProGlyProAlaGlyProProGlyAlaProGly-ProAl 231

QY 369 GCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 319
Db 231 ArgLysProGlyAlaHisGlyProAlaGlyProLysGlyAlaHisGlyProAlaGlyPro 251

QY 318 CAAGCAGCGGCTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 262
Db 251 Olys---GlyAlaHisGlyProAlaGlyProLysGlyAla-HisGlyProAlaGlyProL 270

QY 261 AAGCGCGCGG---CCATGCGAGACCGCGCGCGCGCGCT---CCTCCCGCGCGAGGCG 208
Db 270 YsgLysAlaProGlyProAlaGlyProProGlyAlaProGlyProAlaGlyProPro- 288

QY 207 TTGCTGCGCGCTGCGCGCGCGCGGAGACTCTG-----CGCGCGGTGG 169
Db 289 --GlyAlaProGlyProAlaGlyProProGlyAlaProGlyProAlaGlyProProGlyAl 308

QY 168 CACCGCGCTCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 133
Db 308 LProGlyProAlaGlyProProGlyAlaProGlyProAlaGlyProProGlyAlaProG 328

QY 132 GGCATCAGAGATGCGAGGCGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 73
Db 328 LProAlaGlyProProGlyAlaProGlyProAlaGlyProProGlyAlaProGlyProAl 348

QY 72 CCGCGCTTGGCGCTGCGAGCTCCGAGCGCTGCGCGCGCTCTCAGAGAGAGGTCAGG 13
Db 348 LArgLysProProGlyAlaProGlyProAlaGlyProProGlyAla-----ProG 364

QY 12 GCCCAGAGAGA 2
Db 364 LProAlaGly 367

```

Search completed: February 10, 2004, 17:45:19
Job time : 40.5 secs

Alignment Scores: 3.35e-35 Length: 135
 Pred. No.: 631.50 Matches: 124
 Score: 71.11% Conservative: 7
 Percent Similarity: 68.89% Mismatches: 4
 Best Local Similarity: 45.01% Indels: 45
 Query Match: 12 Gaps: 1
 DB:

US-10-023-182-1 (1-752) x US-10-295-027-388 (1-135)

QY 54 ATGCAGGCCGAGAGCCGCGGCAACAGGGGTTGACGCGGATGCTGATGCGCCAGAGGC 113
 DB 1 MetGlnAlaGlnGlyGlnGlyThrGlySerThrGlyAspAlaAspGlyProGlyGly 20
 QY 114 CCTGGCATTCCTGATGCGCCAGAGGCAATGCTGCGGCGCCAGAGAGCGGGTCCAG 173
 DB 21 ProGlyIleProAspGlyProGlyGlyAlaGlyGlyProGlyGlyAlaThr 40
 QY 174 GCGCGCAGAGTCCCGCGGCGCAGAGGCAAGAGGCTCCGCGCCGCGAGAGCGCC 233
 DB 41 GlyGlyArgGlyProArgGlyAlaGlyAlaAlaArgAlaSerGlyProArgGlyGlyAla 60
 QY 234 CCGCGGGTCCGATGCGCGCGCGCTTCAGAGGCTGAATGCTGAGATGCGAGATGCGGGCC 293
 DB 61 ProArgGlyProHisGlyGlyAlaAlaSerAlaGlnAspGlyArgCysProCysGlyAla 80
 QY 294 AGGAGGCGGAGAGCCGCTGCTTGAATCTTACCTCGCCATGCTTTGCGCAGACCCATG 353
 DB 81 ArgArgProAspSerArgLeuGlnPhe----- 90
 QY 354 GAAGCAGAGCTGCGCCGAGAGCTGCGCCAGAGATGCCACCGCTTCCGTCGAGG 413
 DB 90 ----- 90
 QY 414 GTGCTTGAAGAGTCACTGTGTCCGCAACATAGTATCCGACTGAGTGTGCA 473
 DB 91 -----ArgLeuThrAlaAla 95
 QY 474 GACCAACCGCAATGCAAGCTTCTGATCAGTCTCTGCTCCAGAGTTCCTCTGTGAG 533
 DB 96 AspHisArgGlnLeuGlnLeuSerIleSerSerCysLeuGlnGlnLeuSerLeuLeuMet 115
 QY 534 TGATACGCGAGTGTCTGCGCCGCTTTTGGCTCAGCTCCCTCAGAGGCAAGAGCCG 593
 DB 116 TrpIleThrGlnCysPheLeuProValPheLeuAlaGlnAlaProSerGlyGlnArg 135

RESULT 13
 US-10-157-031-88 ; Sequence 88, Application US/10157031
 ; Publication No. US20030108890A1
 ; GENERAL INFORMATION:

APPLICANT: Baranova, A. V.
 APPLICANT: Yankovsky, N. K.
 APPLICANT: Kozlov, A. P.
 APPLICANT: Lobashev, A. V.
 APPLICANT: Krutovskaya, L. L.
 TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
 FILE REFERENCE: 2760-103
 CURRENT FILING DATE: 2002-05-30
 NUMBER OF SEQ ID NOS: 415
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 88
 LENGTH: 210
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-157-031-88

Alignment Scores: 6.43e-33 Length: 210
 Pred. No.: 598.00 Matches: 121
 Score: 68.12% Conservative: 20

Best Local Similarity: 58.45% Mismatches: 38
 Query Match: 42.62% Indels: 28
 DB: 15 Gaps: 4

US-10-023-182-1 (1-752) x US-10-157-031-88 (1-210)

QY 54 ATGCAGGCCGAGAGCCGCGGCAACAGGGGTTGACGCGGATGCTGATGCGCCAGAGGC 113
 DB 1 MetGlnAlaGlnGlyGlnGlyThrGlySerThrGlyAspAlaAspGlyProGlyGly 20
 QY 114 CCTGGCATTCCTGATGCGCCAGAGGCAATGCTGCGGCGCCAGAGAGCGGGTCCAG 173
 DB 21 ProGlyIleProAspGlyProGlyGlyAlaGlyGlyProGlyGlyAlaThr 40
 QY 174 GCGCGCAGAGTCCCGCGGCGCAGAGGCAAGAGGCTCCGCGCCGCGAGAGCGCC 233
 DB 41 GlyGlyArgGlyProArgGlyAlaGlyAlaAlaArgAlaSerGlyProArgGlyGlyAla 60
 QY 234 CCGCGGGTCCGATGCGCGCGCGCTTCAGAGGCTGAATGCTGAGATGCGAGATGCGGGCC 293
 DB 61 ProArgGlyProHisGlyGlyAlaAlaSerAlaGlnAspGlyArgCysProCysGlyAla 80
 QY 294 AGGAGGCGGAGAGCCGCTGCTTGAATCTTACCTCGCCATGCTTTGCGCAGACCCATG 353
 DB 81 ArgArgProAspSerArgLeuGlnPhe----- 100
 QY 354 GAAGCAGAGCTGCGCCGAGAGCTGCGCCAGAGATGCCACCGCTTCCGTCGAGG 413
 DB 101 GlnAlaGlnLeuValArgArgIleLeuSerArgAspAlaAlaProLeuProArgProGly 120
 QY 414 GTGCTTGAAGAGTCACTGTGTCCGCGCAACATAGTATCCGACTGAGTGTGCA 473
 DB 121 AlaValLeuIleValAspPheThrValSerGlyAsnLeuLeuPheMetSerValArgAspGln 140
 QY 474 GACCAACCGCAATGCAAGCTTCTGATCAGTCTCTGCTCCAGAGTTCCTCTGTGAG 533
 DB 141 AspArgGlnGly-----AlaGlyArgMetArgValGly 152
 QY 534 TGATACGCGAGTGTCTGCGCCGCTTTTGGCTCAGCTCCCTCAGAGGCAAGAGG--- 590
 DB 153 Trp-----GlyLeuGlySerAlaSerProGlnGlyGlnAla 165
 QY 591 -----CGTAAAGCCAGAGCTGCGCCCTTCTCTAG 620
 DB 166 ArgAspLeuArgThrProGlyHisGlyValSerGlnGlnArgProGlyIleProGly--- 184
 QY 621 GTCATGCTCCTCCCTCAGAGG 641
 DB 185 ---ProProProProGlnGly 190

RESULT 14
 US-10-117-937-76 ; Sequence 76, Application US/10117937
 ; Publication No. US20030220239A1
 ; GENERAL INFORMATION:

APPLICANT: CTL IMMUNO THERAPIES CORP.
 APPLICANT: SIMARD, John, J.L.
 APPLICANT: DIAMOND, David, C.
 APPLICANT: Liu, Liping
 APPLICANT: XIE, Zhidong
 TITLE OF INVENTION: EPITOPE SEQUENCES
 FILE REFERENCE: CILIMM.027A
 CURRENT FILING DATE: 2002-04-04
 PRIOR APPLICATION NUMBER: US 60/282,211
 PRIOR FILING DATE: 2001-04-06
 PRIOR APPLICATION NUMBER: US 60/337,017
 PRIOR FILING DATE: 2001-11-07
 PRIOR APPLICATION NUMBER: US 60/363,210
 PRIOR FILING DATE: 2002-03-07
 NUMBER OF SEQ ID NOS: 602
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 76

```
Db 41 G1G1YARGLYPROAGTGLYALAGLYALALALARGALASERGLYPROARGGLYGLYALA 60
QY 234 CCGCGGGGTCCGCATGGCGGGCGGCTTCAGGGCTGAATGATGCTGCAGATCGGGGCC 293
Db 61 PROARGGLYPROHISGLYGLYALALASERLALAGLNPARGCYSPROCYSGLYALA 80
QY 294 AGGGGCGCGAGAGCCCGCTGCTTGAATTCATCCCTCCCATGCTTTGGGACACCCATG 353
Db 81 ARGARGPROASPSERARGLEULEUGLNUHISTLEHMETPROPHESERSETPROMET 100
QY 354 GAACGAGAGCTGGCCCGGAGAGAGCTGGCCCGAGATGCCCAACCGCTTCCCGTCCAGG 413
Db 101 GIUALAGLNULEVALARGARGLLEULEUSERARGSPALALAPROLEUPROARGPROGLY 120
QY 414 GTGCTTGAAGAGATTCACTGTGTCCGCAACATAGTGAATCCGACTGACTGCA 473
Db 121 ALAVALLEULYASAPHETHVALSERGLYASNULEUHEHLEARGLEUTHRALALA 140
QY 474 GACCAACCGCAACTGCAGCTTCATCAGCTCCTGTCTCCAGCAGCTTTCCCTGTTGATG 533
Db 141 ASPHISARGLNULEUGLNULEUSERLLEUSERCYSEUGLNULEUSERLEULEMET 160
QY 534 TGAATCAGCAGAGTCTTCTGCGCGCTGTTTGGCTCAGCCTCCCTCAGGAGAGAGCGC 593
Db 161 TPLETHRGINCYPHELEUPROVALPHELEUALAGLINALPROSERGLYGLINARGARG 180

RESULT 11
US-10-146-473-69
; Sequence 69, Application US/10146473
; Publication No. US20030108888A1
; GENERAL INFORMATION:
; APPLICANT: Scallan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JPV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 69
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-473-69

Alignment Scores:
Pred. No.: 1,19e-47 Length: 180
Score: 814.00 Matches: 152
Percent Similarity: 91.67% Conservative: 13
Best Local Similarity: 84.44% Mismatches: 15
Query Match: 58.02% Indels: 0
Gaps: 0

US-10-023-182-1 (1-752) x US-10-146-473-69 (1-180)
QY 54 ATGACGCGCCGAAGCCCGGGGCGACAGGGGCTTCACGCGGCATGCTGATGCGCCGAGAGCC 113
Db 1 METGLNALAGLNGLNGLYTHRGLYSERTHRGLYSPALASPSGLYPROGLYGLY 20
QY 114 CCTGGCATTCCTGATGCGCCGAGGGGCGCATGCTGCGCGCCGACGAGAGAGCGGCTGCACG 173
Db 21 PROGLYLEPROASPSGLYPROGLYGLYASNLALAGLYGLYPROGLYGLYALALATHR 40
QY 174 GCGCGCAGAGGTCCTCCCGGGGCGCAGGGGCGAGCAAGGCTTCGGGGCGCGGAGAGAGCGCC 233
Db 41 GLYGLYARGGLYPROARGGLYALAGLYALALALARGALASERGLYPROARGGLYGLYALA 60
```

```
QY 234 CCGCGGGGTCCGCATGGCGGGCGGCTTCAGGGCTGAATGATGCTGCAGATCGGGGCC 293
Db 61 PROARGGLYPROHISGLYGLYALALASERLALAGLNPARGCYSPROCYSGLYALA 80
QY 294 AGGGGCGCGAGAGCCCGCTGCTTGAATTCATCCCTCCCATGCTTTGGGACACCCATG 353
Db 81 ARGARGPROASPSERARGLEULEUGLNUHISTLEHMETPROPHESERSETPROMET 100
QY 354 GAACGAGAGCTGGCCCGGAGAGAGCTGGCCCGAGATGCCCAACCGCTTCCCGTCCAGG 413
Db 101 GIUALAGLNULEVALARGARGLLEULEUSERARGSPALALAPROLEUPROARGPROGLY 120
QY 414 GTGCTTGAAGAGATTCACTGTGTCCGCAACATAGTGAATCCGACTGACTGCA 473
Db 121 ALAVALLEULYASAPHETHVALSERGLYASNULEUHEHLEARGLEUTHRALALA 140
QY 474 GACCAACCGCAACTGCAGCTTCATCAGCTCCTGTCTCCAGCAGCTTTCCCTGTTGATG 533
Db 141 ASPHISARGLNULEUGLNULEUSERLLEUSERCYSEUGLNULEUSERLEULEMET 160
QY 534 TGAATCAGCAGAGTCTTCTGCGCGCTGTTTGGCTCAGCCTCCCTCAGGAGAGAGCGC 593
Db 161 TPLETHRGINCYPHELEUPROVALPHELEUALAGLINALPROSERGLYGLINARGARG 180

RESULT 12
US-10-295-027-388
; Sequence 388, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afari, Daniela
; APPLICANT: Afari, Daniela
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevizi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Bos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 388
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-388
```

Db 21 ProGlyIleProAspGlyProGlyGlyAsnAlaGlyIleProGlyGlyAlaGlyAlaThr 40

QY 174 GCGCGCAGAGGTCCCGCGCGCGCAGAGGCGCAGCAAGAGGCTCGCGCGCGCGAGAGGCGCC 233

Db 41 GlyIleArgAlaProArgGlyAlaGlyAlaIleArgAlaSerGlyProGlyGlyAla 60

QY 234 CCGCGCGGTCCGCAATGCGCGCGCGCTTCAGAGGCTGGAATGATCTGTGAAGATGCGCGGCGC 293

Db 61 ProArgGlyProHisGlyGlyAlaAlaSerGlyLeuAsnGlyCysCysArgCysGlyAla 80

QY 294 AGGCGCGCGAGAGCGCGCTGTGATCTAAGCTCCGCAATGCTTTGCGCAGACCCATG 353

Db 81 ArgGlyProGlySerArgIleuLeuGlyIleuAlaMetProPheAlaThrPromet 100

QY 354 GAAGCAGAGCTGGCGCGCAGAGCTGGCCAGATGCCCAAGCTTCCGCGCGAGG 413

Db 101 GluAlaGlyLeuAlaArgArgSerLeuAlaGlnAspAlaProProLeuProValProGly 120

QY 414 GTGCTTCTGAAGAGTTCAGTGTCTGCGCAACATAGTACTGATCCGACTGACTGCA 473

Db 121 ValLeuLeuGlyGlyIleuPheThrValSerGlyAsnIleuThrIleArgLeuThrAlaAla 140

QY 474 GACCAACCGCAACTGCAAGCTTTCATCAGCTCTGTCTTCAGAGCTTTCCCTGTTGATG 533

Db 141 AspHisArgGlnLeuGlnLeuSerIleSerSerCysLeuGlnGlnLeuSerLeuLeuMet 160

QY 534 TGGATCAGCAGAGTCTTCTGCGCGCGTGTGCTGAGCTCCCTCAGGCGCAGAGCGCG 593

Db 161 TrpIleThrGlnCysPheLeuProValPheLeuAlaGlnProProSerGlyGlnArgArg 180

RESULT 9

US-10-364-614-14

Sequence 14, Application US/10364614

Publication No. US20030175250A1

GENERAL INFORMATION:

APPLICANT: JANGER, Elke

APPLICANT: KNUTH, Alexander

APPLICANT: OLD, Lloyd

APPLICANT: Gnjaetic, Sacha

TITLE OF INVENTION: ISOLATED PEPTIDES WHICH BIND TO HLA MOLECULES AND USES THEREOF

FILE REFERENCE: LUD 5726.1 CIP

CURRENT APPLICATION NUMBER: US/10/364,614

PRIOR FILING DATE: 2003-02-24

PRIOR APPLICATION NUMBER: US 60/355,828

PRIOR FILING DATE: 2002-02-13

NUMBER OF SEQ ID NOS: 17

SOFTWARE: PatentIn version 3.2

SEQ ID NO 14

LENGTH: 180

TYPE: PRT

ORGANISM: Homo sapiens

US-10-364-614-14

Alignment Scores:

Pred. No.: 9.6e-55 Length: 180

Score: 918.00 Matches: 174

Percent Similarity: 96.67% Conservative: 0

Best Local Similarity: 96.67% Mismatches: 6

Query Match: 65.43% Indels: 0

Gaps: 0

US-10-023-182-1 (1-752) x US-10-364-614-14 (1-180)

QY 54 ATGCAGGCGCAAGCGCGCGCGCAGAGGCGGTTCAGCGGCGATGCTGATGCGCCAGAGGCG 113

Db 1 MetGlnAlaGlnGlyArgGlyThrGlySerThrGlyAspAlaSerGlyProGlyGly 20

QY 114 CCGCGCAATTCCTGATGCGCCAGAGGCGCAATGCTGCGCGCGCCAGAGAGCGCGGTGCGACG 173

Db 21 ProGlyIleProAspGlyProGlyGlyAsnAlaGlyIleProGlyGlyAlaGlyAlaThr 40

QY 174 GCGCGCAGAGGTCCCGCGCGCGCAGAGGCGCAGCAAGGCGCTCGCGCGCGCGAGAGGCGCGC 233

Db 41 GlyIleArgAlaProArgGlyAlaGlyAlaAlaArgAlaSerGlyProGlyGlyAla 60

QY 234 CCGCGCGGTCCGCAATGCGCGCGCGCTTCAGAGGCTGGAATGATGCTGAGATGCGCGGCGC 293

Db 61 ProArgGlyProHisGlyGlyAlaAlaSerGlyLeuAsnGlyCysCysArgCysGlyAla 80

QY 294 AGGCGCGCGAGAGCGCGCTGTGATCTAAGCTCCGCAATGCTTTGCGCAGACCCATG 353

Db 81 ArgGlyProGlySerArgIleuLeuGlyIleuAlaMetProPheAlaThrPromet 100

QY 354 GAAGCAGAGCTGGCGCGCAGAGCTGGCCAGATGCCCAAGCTTCCGCGCGAGG 413

Db 101 GluAlaGlyLeuAlaArgArgSerLeuAlaGlnAspAlaProProLeuProValProGly 120

QY 414 GTGCTTCTGAAGAGTTCAGTGTCTGCGCAACATAGTACTGATCCGACTGACTGCA 473

Db 121 ValLeuLeuGlyGlyIleuPheThrValSerGlyAsnIleuThrIleArgLeuThrAlaAla 140

QY 474 GACCAACCGCAACTGCAAGCTTTCATCAGCTCTGTCTTCAGAGCTTTCCCTGTTGATG 533

Db 141 AspHisArgGlnLeuGlnLeuSerIleSerSerCysLeuGlnGlnLeuSerLeuLeuMet 160

QY 534 TGGATCAGCAGAGTCTTCTGCGCGCGTGTGCTGAGCTCCCTCAGGCGCAGAGCGCG 593

Db 161 TrpIleThrGlnCysPheLeuProValPheLeuAlaGlnProProSerGlyGlnArgArg 180

RESULT 10

US-10-117-937-75

Sequence 75, Application US/10117937

Publication No. US20030220239A1

GENERAL INFORMATION:

APPLICANT: SIMARD, John, J.L.

APPLICANT: DIAMOND, David, C.

APPLICANT: LIU, Zhihong

APPLICANT: XIE, Zhihong

TITLE OF INVENTION: EPITOP SEQUENCES

FILE REFERENCE: CITIMM 027A

CURRENT APPLICATION NUMBER: US/10/117,937

PRIOR FILING DATE: 2002-04-04

PRIOR APPLICATION NUMBER: US 60/282,211

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: US 60/337,017

PRIOR FILING DATE: 2001-11-07

PRIOR APPLICATION NUMBER: US 60/363,210

PRIOR FILING DATE: 2002-03-07

NUMBER OF SEQ ID NOS: 602

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 75

LENGTH: 180

TYPE: PRT

ORGANISM: Homo sapiens

US-10-117-937-75

Alignment Scores:

Pred. No.: 3.97e-48 Length: 180

Score: 821.00 Matches: 154

Percent Similarity: 91.67% Conservative: 11

Best Local Similarity: 85.56% Mismatches: 15

Query Match: 58.52% Indels: 0

Gaps: 0

US-10-023-182-1 (1-752) x US-10-117-937-75 (1-180)

QY 54 ATGCAGGCGCAAGCGCGCGCGCAGAGGCGGTTCAGCGGCGATGCTGATGCGCCAGAGGCG 113

Db 1 MetGlnAlaGlnGlyArgGlyThrGlySerThrGlyAspAlaSerGlyProGlyGly 20

QY 114 CCGCGCAATTCCTGATGCGCCAGAGGCGCAATGCTGCGCGCGCCAGAGAGCGCGGTGCGACG 173

Db 21 ProGlyIleProAspGlyProGlyGlyAsnAlaGlyIleProGlyGlyAlaGlyAlaThr 40

QY 174 GCGCGCAGAGGTCCCGCGCGCGCAGAGGCGCAGCAAGGCGCTCGCGCGCGCGAGAGAGGCGCGC 233

ପଞ୍ଚମେ ୫

QY	5	ATGAGGCGCAAGACCGGGGCAACAGGAGGGTTGACGGGCGATGCTGATGGCCAGAGAGG	113
Db	1	MetGlnAlaIaIuIuIyArGgIyHrGyIySeTrrGyIyAspAlaAspGlyProGlyGly	20
QY	114	CCTGGCAATTCCTGATGCGCCAGGAGGCAATGCTGGCGGCCAGAGAGGCGGGTCCAG	173
Db	21	ProGlyIyLeProAspGlyProGlyGlyAsnAlaGlyGlyProGlyGluAlaGlyAlaThr	40
QY	174	GCGCGCAGAGGTCCTCCGGGGCGCAGGGGCGACAGAGGCTCGGGGCGGAGAGGCGCC	233
Db	41	GlyGlyArGgIyArProArGgIyAlaGlyAlaAlaArGgIyAspGlyProGlyGlyGlyAla	60
QY	234	CCGGGGGGTCGCGATGGCGCGCGCGGCTTACAGGGCTGAATGATGTGTCAGATGCGGGCC	293
Db	61	ProArGgIyProHrAsGlyGlyAlaAlaAspGlyLeuAsnGlyCyGcYArGcGlyAla	80
QY	294	AGGGGGCGGAGAGCGCGCTGTGAATCTTACTTCGCGCATCGCTTTCGCAACCCATG	353
Db	81	ArgGlyProGluSerArGLeuLeuGluPheTyLeuAlaIeMetProPheAlaThrProMet	100
QY	354	GAGGCAAGTTCGCCCGCGAGAGCTGGCGCAGAGAGCCCGCTTCGGTGCAGGG	413
Db	101	GluAlaGluLeuAlaArGArGSerLeuAlaGlnAspAlaProProLeuProValProGly	120
QY	414	GTCGTTTCGAGAGGTTCACTGTGTCCGGCAATCACTACATATCCGACTGCTTCA	473
Db	121	ValLeuLeuIySgIuPheHrValSerGlyAsnIleLeuThrIleArgLeuHrAlaAla	140
QY	474	GACCAACGCCAATGCAGCTTCCTCATACGCTCTGTCTCCAGACGCTTTCCTGTTATG	533

QY 54 ATGACAGCCGCAAGCCCGGGGACACAGGGGGTTTCAGACGGGCATGCTGTAATGGCCCAAGAGAGC 117
Db 1 MetGlnAlaIleIuIyArGgIyIhTrGyIySerIhTrGyAlaSpAlaAspGlyPrGgIyLy 20
QY 114 CTGGGCACTTCGTAATGAGCCAGGGGGCAATGCTGGGGGCCAGAGAGAGCGGGTGTCCAG 172

```

/ GENERAL INFORMATION:
/ APPLICANT: Ledbetter, Jeffrey A.
/ APPLICANT: Hayden-Ledbetter, Martha S.
/ TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
/ FILE REFERENCE: 390069,401C1
/ CURRENT APPLICATION NUMBER: US/10/207,655
/ CURRENT FILING DATE: 2002-07-25
/ NUMBER OF SEQ ID NOS: 426
/ SOFTWARE: Patent version 3.0
/ SEQ ID NO 71
/ LENGTH: 180
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-207-655-71

Alignment Scores:
Pred. No.: 1,53e-57 Length: 180
Score: 959.00 Matches: 180
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.35% Indels: 0
Gaps: 0
DB: 15

US-10-023-182-1 (1-752) x US-10-207-655-71 (1-180)

QY 54 ATGACAGCCGAGAGCCGAGGCGACAGAGGCGTTCGATGCGCCAGAGAGC 113
DB 1 MetGlnAlaGluGlyArgGlyThrGlySerThrGlyAspAlaAspGlyProGlyGly 20
QY 114 CTTGGCATTCCTGATGCGCCAGAGGCGCAATGCTGCGCGCCAGAGAGCGGTCGACG 173
DB 21 ProGlyIleProAspGlyProGlyGlyAsnAlaGlyGlyProGlyGlnAlaGlyAlaThr 40
QY 174 GCGGCGAGAGGTCCTCCGAGGCGAGAGGCGCAATGCTGCGCGCCAGAGAGCGCGCC 233
DB 41 GlyGlyArgGlyProArgGlyAlaGlyAlaAlaArgAlaSerGlyProGlyGlyGlyAla 60
QY 234 CCGCGGGGTCCGATGCGCGCGCGCTTCAGGGCTGAATGATGCTGAGATGCGGCGCC 293
DB 61 ProArgGlyProIleGlyAlaGlyAlaAlaSerGlyLeuAsnGlyCysCysArgCysGlyAla 80
QY 294 AGGAGGCGCGAGAGCCGCTGCTTGAATCTTACCTCGCATTCCTTTGCGACACCATG 353
DB 81 ArgGlyProGlnSerArgLeuLeuGlnPheTyrlLeuAlaMetProPheAlaThrProMet 100
QY 354 GAAGCAGAGCTGCGCCGAGAGAGCTGCGCCAGAGATGCCACCGCTTCCGTCGACGG 413
DB 101 GlnAlaGlnLeuAlaArgArgSerLeuAlaGlnAspAlaProProLeuProValProGly 120
QY 414 GTGCTTCTGAAGAGTTCACTGTGTCCGCAACATCTGACTATCCGACTGACTGCTGCA 473
DB 121 ValLeuLeuGlyGlnPheThrValSerGlyAsnIleLeuThrIleArgLeuThrAlaAla 140
QY 474 GACCAACCGCACTGAGCTTCATCAGCTCCTGCTCCAGACAGTTTCCCTGTGATG 533
DB 141 AspIleArgGlnLeuGlnLeuSerIleSerSerCysLeuGlnGlnLeuSerLeuMet 160
QY 534 TGGATCAGCAGAGTCTTTCGCGCGGTGTTTGGCTCAGCCTCCCTCAGAGGAGAGCGC 593
DB 161 TrrIleThrGlnCysPheLeuProValPheLeuAlaGlnProProSerGlyGlnArgArg 180

```

```

/ CURRENT FILING DATE: 2001-03-30
/ PRIOR APPLICATION NUMBER: US 60/193,504
/ PRIOR FILING DATE: 2000-03-30
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 27
/ LENGTH: 397
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: NY-ESO-1C tumor antigen
US-09-821-883-27

Alignment Scores:
Pred. No.: 1,53e-57 Length: 397
Score: 959.00 Matches: 180
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.35% Indels: 0
Gaps: 9
DB: 9

US-10-023-182-1 (1-752) x US-09-821-883-27 (1-397)

QY 54 ATGACAGCCGAGAGCCGAGGCGACAGAGGCGTTCGATGCGCCAGAGAGC 113
DB 1 MetGlnAlaGluGlyArgGlyThrGlySerThrGlyAspAlaAspGlyProGlyGly 20
QY 114 CTTGGCATTCCTGATGCGCCAGAGGCGCAATGCTGCGCGCCAGAGAGCGGTCGACG 173
DB 21 ProGlyIleProAspGlyProGlyGlyAsnAlaGlyGlyProGlyGlnAlaGlyAlaThr 40
QY 174 GCGGCGAGAGGTCCTCCGAGGCGAGAGGCGCAATGCTGCGCGCCAGAGAGCGCGCC 233
DB 41 GlyGlyArgGlyProArgGlyAlaGlyAlaAlaArgAlaSerGlyProGlyGlyGlyAla 60
QY 234 CCGCGGGGTCCGATGCGCGCGCGCTTCAGGGCTGAATGATGCTGCGAGATCGGCGGC 293
DB 61 ProArgGlyProIleGlyAlaGlyAlaAlaSerGlyLeuAsnGlyCysCysArgCysGlyAla 80
QY 294 AGGAGGCGCGAGAGCCGCTGCTTGAATCTTACCTCGCATTCCTTCCGACACCATG 353
DB 81 ArgGlyProGlnSerArgLeuLeuGlnPheTyrlLeuAlaMetProPheAlaThrProMet 100
QY 354 GAAGCAGAGCTGCGCCGAGAGAGCTGCGCCAGAGATGCCACCGCTTCCGTCGACGG 413
DB 101 GlnAlaGlnLeuAlaArgArgSerLeuAlaGlnAspAlaProProLeuProValProGly 120
QY 414 GTGCTTCTGAAGAGTTCACTGTGTCCGCAACATCTGACTATCCGACTGACTGCTGCA 473
DB 121 ValLeuLeuGlyGlnPheThrValSerGlyAsnIleLeuThrIleArgLeuThrAlaAla 140
QY 474 GACCAACCGCACTGAGCTTCATCAGCTCCTGCTCCAGACAGTTTCCCTGTGATG 533
DB 141 AspIleArgGlnLeuGlnLeuSerIleSerSerCysLeuGlnGlnLeuSerLeuMet 160
QY 534 TGGATCAGCAGAGTCTTTCGCGCGGTGTTTGGCTCAGCCTCCCTCAGAGGAGAGCGC 593
DB 161 TrrIleThrGlnCysPheLeuProValPheLeuAlaGlnProProSerGlyGlnArgArg 180

```

```

RESULT 6
US-09-821-883-27
/ Sequence 27, Application US/09821883
/ Patent No. US20020061310A1
/ GENERAL INFORMATION:
/ APPLICANT: Laus, Reiner
/ APPLICANT: Vladoic, Damir
/ APPLICANT: Gradiic, Thomas
/ TITLE OF INVENTION: Compositions and Methods for Dendritic
/ FILE REFERENCE: 7636-0022.30
/ CURRENT APPLICATION NUMBER: US/09/821,883

```

```

RESULT 7
US-09-751-798-8
/ Sequence 8, Application US/09751798
/ Patent No. US20020010321A1
/ GENERAL INFORMATION:
/ APPLICANT: Stockert, Elisabeth; Jager, Elke;
/ APPLICANT: Chen, Yao-tseng; Scanlan, Matthew;
/ APPLICANT: Knuth, Alexander; Old, Lloyd J.
/ TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer
/ TITLE OF INVENTION: Associated Proteins, Uses Thereof,
/ TITLE OF INVENTION: Truncated Forms of NY-ESO-1, and HLA
/ TITLE OF INVENTION: Binding Peptides Derived Therefrom
/ NUMBER OF SEQUENCES: 8
/ CORRESPONDENCE ADDRESS:

```

Alignment Scores:

Pred. No.:	1.53e-57	Length:	180
Score:	959.00	Matches:	180
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	68.35%	Indels:	0
DB:	12	Gaps:	0

US-10-023-182-1 (1-752) x US-10-117-937-74 (1-180)

```

QY 54 ATGACAGCCGGAAGCCGCGGAGCAGAGGCGGCTTCAGCGGGCGATGCTGATGCGCCAGAGGC 113
Db 1 MetGlnAaGlnGlyAaGlyThrGlySerThrGlyAspAlaAspGlyProGlyGly 20
QY 114 CTTGCAATTCCTGATGCGCCAGAGGCGCAATGCTGCGCGCCAGAGAGGCGGATGCGCAG 173
Db 21 ProGlyIleProAspGlyProGlyGlyAaAlaAspGlyProGlyGlyAaAlaThr 40
QY 174 GCGGCGAGAGGTCCCGGCGGCGAGAGGCGCAAGAGGCGCTCGGGCGCGGAGAGGCGGC 233
Db 41 GlyIleArgGlyProArgGlyAaGlyAaAlaAspGlyProGlyGlyAaAla 60
QY 234 CCGCGGGGTCCGCAATGCGGCGCGGCTTCAGGCGTGAATGCTGCAAGATCGGCGGC 293
Db 61 ProArgGlyProHisGlyGlyAaAlaAspGlyLeuAsnGlyCysArgCysGlyAa 80
QY 294 AGGCGCGCGAGAGCGCGCTGCTGAGTCTACCTCGCCAGCTTTGCGAGACCGCATG 353
Db 81 ArgGlyProGlySerAlaGlyLeuGlnPheLeuAlaMetProPheAlaThrPromet 100
QY 354 GAAGCAGAGGTGCGCGCGCGAGAGGCTTGCAGAGATGCCACCGCTTCCCGTCCAGAG 413
Db 101 GlnAaGlnLeuAaIaArgSerLeuAlaGlnAspAlaProProLeuProValProGly 120
QY 414 GTGCTTTCGAAGGATTCATCTGCTCCGCGCAACTATGACTATCCGACTGCTGCA 473
Db 121 ValLeuLeuGlyGlnPheThrValSerGlyAsnIleLeuThrIleArgLeuThrAlaAla 140
QY 474 GACCAACCGCACTGCACTCTCCATCAGCTCTCTGCTCCAGAGCTTTCCTGTTGATG 533
Db 141 AspHisArgGlnLeuGlnLeuSerIleSerSerGlyLeuGlnLeuSerLeuLeuMet 160
QY 534 TGGATCAGCGAGTCTTTCGCGCGGTGTTTGGCTCAGCCTCCCTTACGCGCAGAGCGC 593
Db 161 TrpIleThrGlnCysPheLeuProValPheLeuAlaGlnProSerGlyGlnArgArg 180

```

RESULT 4

US-10-295-027-386

Sequence 386, Application US/10295027

Publication No. US20030232350A1

GENERAL INFORMATION:

APPLICANT: Afar, Daniel

APPLICANT: Aziz, Natasha

APPLICANT: Ginsberg, Wendy M.

APPLICANT: Gish, Kurt C.

APPLICANT: Glynn, Richard

APPLICANT: Heyez, Peter A.

APPLICANT: Mack, David H.

APPLICANT: Murray, Richard

APPLICANT: Watson, Susan R.

APPLICANT: Eos Biotechnology, Inc.

TITLE OF INVENTION: Methods of diagnosis of Cancer, Compositions and

TITLE OF INVENTION: Methods of screening for Modulators of Cancer

FILE REFERENCE: 018501-012500US

CURRENT APPLICATION NUMBER: US/10/295,027

CURRENT FILING DATE: 2002-11-13

PRIOR APPLICATION NUMBER: US 09/663,733

PRIOR FILING DATE: 2000-09-15

PRIOR APPLICATION NUMBER: US 60/350,666

PRIOR FILING DATE: 2001-11-13

PRIOR APPLICATION NUMBER: US 60/335,394

PRIOR FILING DATE: 2001-11-15

PRIOR APPLICATION NUMBER: US 60/332,464

PRIOR FILING DATE: 2001-11-21

PRIOR APPLICATION NUMBER: US 60/334,393

PRIOR FILING DATE: 2001-11-29

PRIOR APPLICATION NUMBER: US 60/340,376

PRIOR FILING DATE: 2001-12-14

PRIOR APPLICATION NUMBER: US 60/347,211

PRIOR FILING DATE: 2002-01-08

PRIOR APPLICATION NUMBER: US 60/347,349

PRIOR FILING DATE: 2002-01-10

PRIOR APPLICATION NUMBER: US 60/355,250

PRIOR FILING DATE: 2002-02-08

PRIOR APPLICATION NUMBER: US 60/356,714

PRIOR FILING DATE: 2002-02-13

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 1386

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 386

LENGTH: 180

TYPE: PRT

ORGANISM: Homo sapiens

US-10-295-027-386

Alignment Scores:

Pred. No.:	1.53e-57	Length:	180
Score:	959.00	Matches:	180
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	68.35%	Indels:	0
DB:	12	Gaps:	0

US-10-023-182-1 (1-752) x US-10-295-027-386 (1-180)

```

QY 54 ATGACAGCCGGAAGCCGCGGAGCAGAGGCGGCTTCAGCGGGCGATGCTGATGCGCCAGAGGC 113
Db 1 MetGlnAaGlnGlyAaGlyThrGlySerThrGlyAspAlaAspGlyProGlyGly 20
QY 114 CTTGCAATTCCTGATGCGCCAGAGGCGCAATGCTGCGCGCCAGAGAGGCGGATGCGCAG 173
Db 21 ProGlyIleProAspGlyProGlyGlyAaAlaAspGlyProGlyGlyAaAlaThr 40
QY 174 GCGGCGAGAGGTCCCGGCGGCGAGAGGCGCAAGAGGCGCTCGGGCGCGGAGAGGCGGC 233
Db 41 GlyIleArgGlyProArgGlyAaGlyAaAlaAspGlyProGlyGlyAaAla 60
QY 234 CCGCGGGGTCCGCAATGCGGCGCGGCTTCAGGCGTGAATGCTGCAAGATCGGCGGC 293
Db 61 ProArgGlyProHisGlyGlyAaAlaAspGlyLeuAsnGlyCysArgCysGlyAa 80
QY 294 AGGCGCGCGAGAGCGCGCTGCTGAGTCTACCTCGCCAGCTTTCGCGCAACCGCATG 353
Db 81 ArgGlyProGlySerAlaGlyLeuGlnPheThrValSerGlyAsnIleLeuThrIleArgLeuThrAlaAla 100
QY 354 GAAGCAGAGGTGCGCGCGCGAGAGGCTTGCAGAGATGCCACCGCTTCCCGTCCAGAG 413
Db 101 GlnAaGlnLeuAaIaArgSerLeuAlaGlnAspAlaProProLeuProValProGly 120
QY 414 GTGCTTTCGAAGGATTCATCTGCTCCGCGCAACTATGACTATCCGACTGCTGCA 473
Db 121 ValLeuLeuGlyGlnPheThrValSerGlyAsnIleLeuThrIleArgLeuThrAlaAla 140
QY 474 GACCAACCGCACTGCACTCTCCATCAGCTCTCTGCTCCAGAGCTTTCCTGTTGATG 533
Db 141 AspHisArgGlnLeuGlnLeuSerIleSerSerGlyLeuGlnLeuSerLeuLeuMet 160
QY 534 TGGATCAGCGAGTCTTTCGCGCGGTGTTTGGCTCAGCCTCCCTTACGCGCAGAGCGC 593
Db 161 TrpIleThrGlnCysPheLeuProValPheLeuAlaGlnProSerGlyGlnArgArg 180

```

RESULT 5

US-10-207-655-71

Sequence 71, Application US/10207655

Publication No. US2003018592A1

Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 68.35%
 DB: 12

Conservative: 0
 Mismatches: 0
 Indels: 0
 Gaps: 0

US-10-023-182-1 (1-752) x US-09-849-602-30 (1-180)

```

QY 54 ATGCAGCCGGAAGCCCGGAGCAAGAGGAGTTGACGAGGCGATGCTGATGCGCCAGAGGC 113
DB 1 MetGlnaIaGluGlyAArgGlyThrGlyGlySerThrGlyAspAlaAspGlyProGlyGly 20
QY 114 CCGGATTCCTGATGAGCCGAGGAGCAATGCTGCGGCGCCAGAGAGAGGCGGAGCAG 173
DB 21 ProGlyIleProAspGlyProGlyGlyAsnAlaGlyGlyProGlyGluAlaGlyAlaThr 40
QY 174 GCGGCGAGAGGTCCCGGAGCGAGAGGAGCAAGAGGCGCTCGAGGCGCGAGAGAGCGCC 233
DB 41 GlyIyAArgGlyProAArgGlyAlaGlyAlaAlaArgAlaSerGlyProGlyGlyAla 60
QY 234 CCGCGGAGTCCGATGAGCGGCGCGGCTTCAGAGGCTGATGATGCTGCAATGCGGAGC 293
DB 61 ProAArgGlyProHlaGlyAlaAlaSerGlyLeuAanglyCySAArgGlyAla 80
QY 294 AGGAGGCGGAGAGCGCGCTGCTGATGCTTACCTGCGCATGCGCTTCGCGAGACCCATG 353
DB 81 ArgGlyProGluSerArgLeuLeuGluPheThyIleuAlaMetProPheAlaThrProMet 100
QY 354 GAAGCAGAGCTGCGCCGAGAGGCTGCGCCAGATGCCCGCCAGCTTCCTGCGAGG 413
DB 101 GluAlaGluLeuAlaArgArgSerLeuAlaGlnAspAlaProProLeuProValProGly 120
QY 414 GCGCTTCGAGAGAGTCACTGTGCGGCAACATGACTGACTGACTGACTGACTGCA 473
DB 121 ValLeuLeuGluGluPheThrValSerGlyAsnIleLeuThrIleArgLeuThrAlaAla 140
QY 474 GACCAACCGCAACTGAGCTTCATCAGCTCTGCTTCACAGAGCTTCCTGTTGATG 533
DB 141 AspHisArgGlnLeuGlnLeuSerIleSerSerCysLeuGlnGlnLeuSerLeuLeuMet 160
QY 534 TGGATCAGCAGAGCTTCCTGCGCGGTTTGGCTCAGCGCCCTCAGAGGAGAGCGCC 593
DB 161 TrpIleThrGlnCysPheLeuProValPheLeuAlaGlnProProSerGlyGlnArgArg 180

```

RESULT 2
 US-10-026-066-3

Sequence 3, Application US/10026066
 Publication No. US20030215425A1

GENERAL INFORMATION:
 APPLICANT: Simard, John J. L.
 APPLICANT: Diamond, David C.
 TITLE OF INVENTION: EPITOPE SYNCHRONIZATION IN ANTIGEN
 FILE REFERENCE: CTILMM.21C1C
 CURRENT APPLICATION NUMBER: US/10/026,066
 CURRENT FILING DATE: 2001-12-07
 PRIOR APPLICATION NUMBER: 09/561,074
 PRIOR FILING DATE: 2000-04-28
 PRIOR APPLICATION NUMBER: 09/560,465
 PRIOR FILING DATE: 2000-04-28
 PRIOR APPLICATION NUMBER: 09/561,572
 PRIOR FILING DATE: 2000-04-28
 PRIOR APPLICATION NUMBER: 09/561,571
 PRIOR FILING DATE: 2000-04-28
 PRIOR APPLICATION NUMBER: FCI/US01/13806
 PRIOR FILING DATE: 2001-04-27
 NUMBER OF SEQ ID NOS: 89
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 3
 LENGTH: 180
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-026-066-3

Alignment Scores:
 Pred. No.: 1,536-57
 Score: 959.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 68.35%
 DB: 12

Length: 180
 Matches: 180
 Conservative: 0
 Mismatches: 0
 Indels: 0
 Gaps: 0

US-10-023-182-1 (1-752) x US-10-026-066-3 (1-180)

```

QY 54 ATGCAGCCGGAAGCCCGGAGCAAGAGGAGTTGACGAGGCGATGCTGATGCGCCAGAGGC 113
DB 1 MetGlnaIaGluGlyAArgGlyThrGlyGlySerThrGlyAspAlaAspGlyProGlyGly 20
QY 114 CCGGATTCCTGATGAGCCGAGGAGCAATGCTGCGGCGCCAGAGAGAGGCGGAGCAG 173
DB 21 ProGlyIleProAspGlyProGlyGlyAsnAlaGlyGlyProGlyGluAlaGlyAlaThr 40
QY 174 GCGGCGAGAGGTCCCGGAGCGAGAGGAGCAAGAGGCGCTCGAGGCGCGAGAGAGCGCC 233
DB 41 GlyIyAArgGlyProAArgGlyAlaGlyAlaAlaArgAlaSerGlyProGlyGlyAla 60
QY 234 CCGCGGAGTCCGATGAGCGGCGCGGCTTCAGAGGCTGATGATGCTGCAATGCGGAGC 293
DB 61 ProAArgGlyProHlaGlyAlaAlaSerGlyLeuAanglyCySAArgGlyAla 80
QY 294 AGGAGGCGGAGAGCGCGCTGCTGATGCTTACCTGCGCATGCGCTTCGCGAGACCCATG 353
DB 81 ArgGlyProGluSerArgLeuLeuGluPheThyIleuAlaMetProPheAlaThrProMet 100
QY 354 GAAGCAGAGCTGCGCCGAGAGGCTGCGCCAGATGCCCGCCAGCTTCCTGCGAGG 413
DB 101 GluAlaGluLeuAlaArgArgSerLeuAlaGlnAspAlaProProLeuProValProGly 120
QY 414 GCGCTTCGAGAGAGTCACTGTGCGGCAACATGACTGACTGACTGACTGACTGCA 473
DB 121 ValLeuLeuGluGluPheThrValSerGlyAsnIleLeuThrIleArgLeuThrAlaAla 140
QY 474 GACCAACCGCAACTGAGCTTCATCAGCTCTGCTTCACAGAGCTTCCTGTTGATG 533
DB 141 AspHisArgGlnLeuGlnLeuSerIleSerSerCysLeuGlnGlnLeuSerLeuLeuMet 160
QY 534 TGGATCAGCAGAGCTTCCTGCGCGGTTTGGCTCAGCGCCCTCAGAGGAGAGCGCC 593
DB 161 TrpIleThrGlnCysPheLeuProValPheLeuAlaGlnProProSerGlyGlnArgArg 180

```

RESULT 3
 US-10-117-937-74

Sequence 74, Application US/10117937
 Publication No. US20030220239A1

GENERAL INFORMATION:
 APPLICANT: CTL IMMUNO THERAPIES CORP.
 APPLICANT: SIMARD, John, J.L.
 APPLICANT: DIAMOND, David, C.
 APPLICANT: Liu, Liping
 APPLICANT: Xie, Zhidong
 TITLE OF INVENTION: EPITOPE SEQUENCES
 FILE REFERENCE: CTILMM.027A
 CURRENT APPLICATION NUMBER: US/10/117,937
 CURRENT FILING DATE: 2002-04-04
 PRIOR APPLICATION NUMBER: US 60/282,211
 PRIOR FILING DATE: 2001-04-06
 PRIOR APPLICATION NUMBER: US 60/337,017
 PRIOR FILING DATE: 2001-11-07
 PRIOR APPLICATION NUMBER: US 60/363,210
 PRIOR FILING DATE: 2002-03-07
 NUMBER OF SEQ ID NOS: 602
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 74
 LENGTH: 180
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-117-937-74

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: February 10, 2004, 17:34:42 ; Search time 36.5 Seconds
(without alignments)

8627.697 Million cell updates/sec

Title: US-10-023-182-1

Perfect score: 1403

* Sequence: 1 ATCTCTGTCGCGCCCTGACCT.....TAAACTGAGCTACGAAAA 752

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5	
Delop 6.0 , Delext 7.0	

Searched: 801455 segs, 209382283 residues

Total number of hits satisfying chosen parameters: 1602910

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=x1p
-Q=/cg2_1/USPTO.spool_p/US10023182/runat_10022004.171431_3992/app_query.fasta_1.903
-DB=Published Applications AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPC=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdt -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pico -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10023182 @CGN_1.1.13 @runat_10022004.171431_3992
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGPOP=6 -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications AA:

1:	/cg2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
2:	/cg2_6/ptodata/1/pubppa/US06_PUBCOMB.pep.*
3:	/cg2_6/ptodata/1/pubppa/US06_PUBCOMB.pep.*
4:	/cg2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
5:	/cg2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
6:	/cg2_6/ptodata/1/pubppa/US08_PUBCOMB.pep.*
7:	/cg2_6/ptodata/1/pubppa/US08_PUBCOMB.pep.*
8:	/cg2_6/ptodata/1/pubppa/US08_PUBCOMB.pep.*
9:	/cg2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
10:	/cg2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
11:	/cg2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
12:	/cg2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
13:	/cg2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
14:	/cg2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
15:	/cg2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
16:	/cg2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
17:	/cg2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
18:	/cg2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	959	68.4	180	12	US-09-849-602-30	Sequence 30, Appl
2	959	68.4	180	12	US-10-026-066-3	Sequence 3, Appl
3	959	68.4	180	12	US-10-117-937-74	Sequence 74, Appl
4	959	68.4	180	12	US-10-295-027-386	Sequence 186, Appl
5	959	68.4	180	15	US-10-207-655-71	Sequence 71, Appl
6	959	68.4	397	9	US-09-821-983-27	Sequence 27, Appl
7	953	67.9	180	14	US-09-751-798-8	Sequence 8, Appl
8	953	67.9	180	14	US-10-023-182-8	Sequence 8, Appl
9	918	65.4	180	12	US-10-364-614-14	Sequence 14, Appl
10	921	58.5	180	12	US-10-117-937-75	Sequence 75, Appl
11	814	58.0	180	15	US-10-146-473-69	Sequence 69, Appl
12	631.5	45.0	135	12	US-10-295-027-388	Sequence 388, Appl
13	598	42.6	210	15	US-10-157-831-88	Sequence 88, Appl
14	597	42.6	210	12	US-10-117-937-76	Sequence 76, Appl
15	202.5	14.5	504	12	US-10-342-331-5	Sequence 3, Appl
16	202.5	14.5	720	12	US-10-342-331-4	Sequence 4, Appl
17	196	14.0	960	12	US-10-342-331-5	Sequence 5, Appl
18	188	13.4	822	12	US-10-342-331-49	Sequence 49, Appl
19	185.5	13.3	520	12	US-10-168-097A-36	Sequence 36, Appl
20	185.5	13.3	520	12	US-10-239-431A-34	Sequence 34, Appl
21	185.5	13.3	1341	15	US-10-058-124-18	Sequence 18, Appl
22	185.5	13.3	1464	12	US-10-301-822-28	Sequence 28, Appl
23	185.5	13.3	1464	12	US-10-291-265-243	Sequence 243, Appl
24	185.5	13.3	1464	15	US-10-060-036-159	Sequence 159, Appl
25	185.5	13.3	1464	15	US-10-171-311-36	Sequence 36, Appl
26	185.5	13.3	1464	15	US-10-216-705-21	Sequence 21, Appl
27	185.5	13.3	1464	15	US-10-149-352-2	Sequence 2, Appl
28	185.5	13.3	1464	15	US-10-177-293-65	Sequence 65, Appl
29	184.5	13.2	903	15	US-10-156-761-11093	Sequence 11093, A
30	182.5	13.1	1463	12	US-10-402-089-2	Sequence 2, Appl
31	182.5	13.1	1463	12	US-10-402-072A-2	Sequence 2, Appl
32	181	12.9	1366	12	US-10-402-089-10	Sequence 10, Appl
33	181	12.9	1366	12	US-10-402-072A-10	Sequence 10, Appl
34	179	12.8	1449	12	US-10-402-089-8	Sequence 8, Appl
35	179	12.8	1449	12	US-10-402-072A-8	Sequence 8, Appl
36	179	12.8	1466	12	US-10-402-089-4	Sequence 4, Appl
37	179	12.8	1466	12	US-10-402-089-6	Sequence 6, Appl
38	179	12.8	1466	12	US-10-402-072A-4	Sequence 4, Appl
39	179	12.8	1466	12	US-10-402-072A-6	Sequence 6, Appl
40	178.5	12.8	1497	15	US-10-157-031-128	Sequence 128, Appl
41	176	12.6	417	15	US-10-096-986-104	Sequence 104, Appl
42	176	12.6	837	15	US-10-096-986-103	Sequence 103, Appl
43	175.5	12.5	1466	12	US-10-402-089-12	Sequence 12, Appl
44	175.5	12.5	1466	12	US-10-402-072A-12	Sequence 12, Appl
45	175	12.5	274	9	US-09-850-887-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-849-602-30
; Sequence 30, Application US/09849602
; Publication No. US20030165834A1
GENERAL INFORMATION:
; APPLICANT: Scantlan, Matthew J.
; APPLICANT: Old, Lloyd J.
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Chen, Yao-Tseng
; TITLE OF INVENTION: Colon Cancer Antigen Panel
; FILE REFERENCE: L0461/7105(JRV)
; CURRENT APPLICATION NUMBER: US/09/849,602
; CURRENT FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 30
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-849-602-30

Alignment Scores:
Pred. No.: 1.53e-57 Length: 180
Score: 959.00 Matches: 180

PF 18-APR-2002; 2002WO-US12476.
 XX 18-APR-2001; 2001US-284770P.
 PR 10-MAY-2001; 2001US-290492P.
 PR 09-NOV-2001; 2001US-339245P.
 PR 13-NOV-2001; 2001US-350666P.
 PR 29-NOV-2001; 2001US-334370P.
 PR 12-APR-2002; 2002US-372246P.
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Aziz N, Murray R;
 XX
 DR WPI: 2003-093161/08.
 DR N-PSDB; ABX76423.
 DR
 XX
 PT Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer -
 XX
 PS Claim 27; Page 414; 453pp; English.
 XX
 CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridises
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung
 CC cancer-associated polynucleotides and polypeptides are used for
 CC identifying a compound that modulates a lung cancer-associated
 CC polypeptide, for inhibiting proliferation of a lung cancer-associated
 CC cell to treat lung cancer in a patient and for treating a mammal having
 CC lung cancer by administering a modulatory compound identified. The
 CC methods are useful for treating lung cancer, such as small cell lung
 CC cancer, non-small cell lung cancer or other benign or precancerous
 CC lesions, e.g. atelectasis; emphysema, bronchitis, chronic obstructive
 CC pulmonary disease, fibrosis, hyperresponsivity pneumonitis, interstitial
 CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides
 CC and polypeptides are useful for diagnostic purposes and as targets for
 CC screening for therapeutic compounds that modulate lung cancer, such as
 CC antibodies. Sequences ABUS6408-ABUS6745 represent lung cancer-associated
 CC polypeptides of the invention.
 CC
 XX
 SQ Sequence 180 AA;
 Alignment Scores:
 Pred. No.: 3,71e-66 Length: 180
 Score: 959.00 Matches: 180
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 68.35% Indels: 0
 DB: 24 Gaps: 0
 US-10-023-182-1 (1-752) x ABUS6694 (1-180)
 QY 54 ATGACAGCCGAGAGCCGAGGCGACAGAGGAGTTTCAGAGCGATGCTGATGCGCCAGAGAGC 113
 DB 1 MetGlnAlaGlnGlyArgGlyThrGlySerThrGlyAspAlaAspGlyProGlyGly 20
 QY 114 CTTGCAATTCCTGATGCGCCAGAGGCGCAATGCTGCGGCCAGAGAGAGCGGCTGCCACG 173
 DB 21 ProGlyLeuProAspGlyProGlyGlyAsnAlaGlyGlyProGlyGlyAlaThr 40
 QY 174 GCGCGAGAGGTCCTCCGCGGCGAGGCGGAGAGGCGCTCGGCGCGCGGAGAGGCGCC 233
 DB 41 GluGlyArgGlyProArgGlyAlaGlyAlaAlaArgAlaSerGlyProGlyGlyAla 60
 QY 234 CCGCGGGGTCGCGAGCGGCGCGCTTCAAGGCTGATGATGCTGACAGATCGGAGCC 293
 DB 61 ProArgGlyProHisGlyGlyAlaAlaSerGlyLeuAsnGlyCysCysArgCysGlyAla 80
 QY 294 AGGGGGCGGAGAGCGCGCTGAGTTGTAAGTCTGACCTGCGCATGCTTCCGAGACCCATG 353

DB 81 ArgGlyProGlnSerArgLeuLeuGlnPheTyrLeuAlaMetProPheAlaThrProMet 100
 QY 354 GAGCAGAGCTGGCCCGAGAGAGCTGCGCCAGATGCCCCAGCGCTTCCGTGCGAGG 413
 DB 101 GluAlaGlnLeuAlaArgArgSerLeuAlaGlnAspAlaProProLeuProValProGly 120
 QY 414 GTGCTTTCGAGAGATTACCTGTGTCGCGCAACATACGACTTANCCGACTGACGCTCA 473
 DB 121 ValLeuLeuGlnPheThrValSerGlyAsnLeuThrIleArgLeuThrAla 140
 QY 474 GACCAAGCCCACTGACGCTTCAATGAGCTTCTGTCCAGACAGCTTCCCTGTGATG 533
 DB 141 AspHisArgGlnLeuGlnLeuSerIleSerSerCysLeuGlnGlnLeuSerLeuMet 160
 QY 534 TGGATCAGCGAGTGTCTTGTGCGCGGTGTTTGGCTCAGCTCCTCAGGGCAGAGGCGC 593
 DB 161 TrpIleThrGlnCysPheLeuProValPheLeuAlaGlnProProSerGlyGlnArg 180

Search completed: February 10, 2004, 17:33:51
 Job time : 45.5 secs

Db 121 ValLeuLeuYgSLuPheTrIvaLSerGLyASnIleuThrIleArgLeuThraAla 140
 QY 474 GACACCGCCAGCTGAGCTCTCCATCAGCTCCGTCTCCAGAGAGCTTCCCTGTGATG 533
 Db 141 AsphIaArgInLeuGInLeuSerIleSerCysLeuInGInLeuSerLeuMet 160
 QY 534 TGGATCAGCAGAGTCTTTCGCCGCTGTTTGGCTCAGCTCCCTCAGGCGCAGAGCGC 593
 Db 161 TrpIleThrGInCysPheLeuProValPheLeuIaGInProProSerGlyGInArgArg 180
 RESULT 14
 ABUS6508
 ID ABUS6508 standard; Protein; 180 AA.
 AC ABUS6508;
 XX
 DT 02-APR-2003 (first entry)
 XX
 DE Lung cancer-associated polypeptide #101.
 XX
 KM Lung cancer-associated polypeptide; cytostatic; emphysema;
 KM antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KM small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KM chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KM interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX
 OS Unidentified.
 XX
 PN WO200286443-A2.
 XX
 PD 31-OCT-2002.
 XX
 PF 18-APR-2002; 2002WO-US12476.
 XX
 PR 18-APR-2001; 2001US-284770P.
 PR 10-MAY-2001; 2001US-290492P.
 PR 09-NOV-2001; 2001US-339245P.
 PR 13-NOV-2001; 2001US-350666P.
 PR 29-NOV-2001; 2001US-334370P.
 PR 12-APR-2002; 2002US-372246P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Aziz N. Murray R;
 XX
 DR MPI; 2003-093161/08.
 DR N-PsDB; ABX76232.
 XX
 PT Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer -
 XX
 PS Claim 27; Page 270; 453pp; English.
 XX
 CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung
 CC cancer-associated polynucleotides and polypeptides are used for
 CC identifying a compound that modulates a lung cancer-associated
 CC polypeptide, for inhibiting proliferation of a lung cancer-associated
 CC cell to treat lung cancer in a patient and for treating a mammal having
 CC lung cancer by administering a modulatory compound identified. The
 CC methods are useful for treating lung cancer, such as small cell lung
 CC cancer, non-small cell lung cancer or other benign or precancerous
 CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive
 CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial
 CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides
 CC and polypeptides are useful for diagnostic purposes and as targets for
 CC screening for therapeutic compounds that modulate lung cancer, such as

CC antibodies. Sequences ABUS6408-ABUS6745 represent lung cancer-associated
 CC polypeptides of the invention.
 XX
 SQ Sequence 180 AA;
 Alignment Scores:
 Pred. No.: 3,71e-66 Length: 180
 Score: 959.00 Matches: 180
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 68.35% Indels: 0
 DB: 24 Gaps: 0
 US-10-023-182-1 (1-752) x ABUS6508 (1-180)
 QY 54 ATGCAGCGCCGAAAGCCGCGGCGACAGAGGGGCTTCGACGGGCGATGCTGATGGCCGAGAGGC 113
 Db 1 MetGInIaGInGlyArgGlyThrGlyCysThrGlyAspAlaAspGlyProGlyGly 20
 QY 114 CCTGGCATTCTGATGAGCCGAGGCGCAATGCTGCGCGCCGAGAGAGCGGAGTGCAGC 173
 Db 21 ProGlyIleProAspGlyProGlyGlyAsnIaGlyProGlyGlyIaGlyAlaThr 40
 QY 174 GCGGCAAGAGTCCCGGCGCGAGGCGACAGAGGCTCGGGGCGGAGAGGGGCC 233
 Db 41 GlyGlyArgGlyProArgGlyAlaGlyAlaIaArgAlaSerGlyProGlyGlyAla 60
 QY 234 CCGCGGGGTCGCGATGAGCGCGCGGCTTCAGGCGTGAATGATGCTGAGTGGCGGCC 293
 Db 61 ProArgGlyProHisGlyGlyAlaAlaSerGlyLeuAsnGlyCysArgCysGlyAla 80
 QY 294 AGGGGGCGGAGAGCGCGCTTGAGTTCTTACTCTGCGCATGCTTTCGCAACCCATG 353
 Db 81 ArgGlyProGlyIleSerArgLeuLeuGInPheThrIleuAlaMetProPheAlaThrPromet 100
 QY 354 GAAGAGAGCTGGCGCGCGAGAGAGCTGGCCGAGAGTGGCCGCGCTCCGTCGCAAGG 413
 Db 101 GluAlaGInLeuAlaArgArgSerLeuAlaGInPheAlaProProLeuProValProGly 120
 QY 414 GTGCTTTCGAAAGAGTCACTGTGTCCGCAACATACTGACTATTCGACTGACTGCTGCA 473
 Db 121 ValLeuLeuYgSLuPheThrValSerGlyAsnIleuThrIleArgLeuThraAla 140
 QY 474 GACACCGCCAGCTGAGCTCTCCATCAGCTCCGTCTCCAGAGAGCTTCCCTGTGATG 533
 Db 141 AsphIaArgInLeuGInLeuSerIleSerCysLeuInGInLeuSerLeuMet 160
 QY 534 TGGATCAGCAGAGTCTTTCGCCGCTGTTTGGCTCAGCTCCCTCAGGCGCAGAGCGC 593
 Db 161 TrpIleThrGInCysPheLeuProValPheLeuIaGInProProSerGlyGInArgArg 180
 RESULT 15
 ABUS6694
 ID ABUS6694 standard; Protein; 180 AA.
 AC ABUS6694;
 XX
 DT 02-APR-2003 (first entry)
 XX
 DE Lung cancer-associated polypeptide #287.
 XX
 KM Lung cancer-associated polypeptide; cytostatic; emphysema;
 KM antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KM small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KM chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KM interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX
 OS Unidentified.
 XX
 PN WO200286443-A2.
 XX
 PD 31-OCT-2002.
 XX

The invention relates to a new synthetic polypeptide (I) comprising several different segments of at least one parent polypeptide linked together in a different relationship relative to their linkage in the parent polypeptide to impede, sabotage or otherwise alter at least one function associated with the parent polypeptide and for inducing an immune response against a pathogen or cancer. Also included are a synthetic polynucleotide encoding and a computer system for designing the synthetic polypeptides. The synthetic polypeptides and polynucleotides are referred to as a Savane. The synthetic polypeptide is useful for modulating immune responses preferably directed against a pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone marrow, testis, oesophagus, brain, testicle, uterus), as potentiating agents. Compositions comprising the polypeptide may be used in the treatment or prophylaxis against viral (such as infections caused by HIV (human immunodeficiency virus), hepatitis, influenza, Japanese encephalitis virus, Epstein-Barr virus and respiratory syncytial virus), bacterial (e.g., infections caused by *Neisseria meningococcal*, *Haemophilus Salmonella*, *Streptococcal*, *Legionella* and *Mycobacterium* or parasitic (e.g., infections caused by *Plasmodium*, *Schistosoma*, *Leishmania*, *Trypanosoma*, *Toxoplasma* and *Giardia*) infections. The present sequence is a consensus sequence for a parent protein used to design a Savane of the invention.

US-10-023-182-1 (1-752) x AAU84818 (1-180)

[illegible]

RESULT 12
ID AAVU1543 standard; protein; 180 AA.
AAU1543
XX AAVU1543;
XX
DT 12-MAR-2002 (first entry)
XX
XX Human tumour associated antigen NY-ESO.
DE
XX Human; tumour associated antigen; NY-ESO; human leukocyte antigen;
KW major histocompatibility complex; MHC; HLA-A2.2; vaccine; cancer;
KW HIV; human immunodeficiency virus infection; cytostatic; virucide;
KW housekeeping epitope; adoptive immunotherapy; neoplastic disease;
KW viral diseases; hepatitis virus; papilloma virus; tumour; leukaemia;
KW lymphoma; breast cancer; prostate cancer; lung cancer;
KW parasitic infection; Chlamydia; Trypanosoma; Toxoplasma.
XX
XX Homo sapiens.
OS
PN WO200182963-A2.
XX
XX
PD 08-NOV-2001.
XX
XX 27-APR-2001; 2001WO-US13806.
PE
XX 28-APR-2000; 2000US-0560465.
FR 28-APR-2000; 2000US-0561074.
PR 28-APR-2000; 2000US-0561571.
PR 28-APR-2000; 2000US-0561572.
XX
XX
PA (CTL1-) CTL IMMUNOTHERAPIES CORP.
XX
PI Simard JUI, Diamond DC, Lei X;
XX
DR WPI: 2002-066492/09.
XX
PT Novel vaccine useful for treating neoplastic and viral diseases;
PT comprises a first housekeeping epitope derived from a first antigen
PT associated with a first target cell -
XX
XX Example 23; Fig 17; 131pp; English.
XX
XX The invention relates to a vaccine comprising a first housekeeping
CC epitope derived from a first antigen associated with a first target
CC cell. Also included are an isolated T cell expressing a T cell receptor
CC specific for a major histocompatibility complex (MHC)-peptide complex
CC comprising a first housekeeping epitope which is derived from a first
CC antigen associated with a first target cell, selecting an epitope
CC (or peptide sequence) from a population of peptide fragments of an
CC antigen associated with a target in a host, where the fragments have a
CC known or predicted affinity for a MHC receptor peptide binding cleft of
CC the host, where the epitope selected corresponds to a product of
CC proteolytic cleavage of the antigen in a cell of the host and
CC a nucleic acid construct comprising a first coding region, where the
CC first coding region comprises a first sequence encoding at least a first
CC polypeptide, where the first polypeptide comprises a first target cell;
CC epitope derived from a first antigen associated with a first target cell;
CC The epitopes, peptides, vaccines and nucleic acids are useful in the
CC manufacture of a medicament for use in adoptive immunotherapy and for
CC prevention and treatment of neoplastic and viral diseases (e.g.
CC human immunodeficiency virus, HIV, infection, hepatitis virus and
CC papilloma virus), cancers (e.g. tumours, leukaemia, lymphoma, breast
CC cancer, prostate cancer and lung cancer), infection of cells by
CC intracellular parasites (e.g. Chlamydia, Trypanosoma and
CC Toxoplasma) and many other examples given in the specification.
CC The invention permits the vaccine designer to ignore peptides that,
CC despite predicted high binding affinity for MHC, will never be useful
CC because they cannot be presented by target cells. The invention provides
CC a major advance in vaccine design, one that combines the power of antigen
CC sequence analysis with the fundamental realities of immunology. The
CC invention allows for the simple and effective selection of meaningful

QY 474 GACCAACCGCAACTGCAGCTCTCCATCAGCTCTGTCTCCAGAGCTTTCCCTGTATG 533
 DB 141 AspHisArgGlnLeuGlnLeuSerIleSerSerCysLeuGlnLeuSerIleuMet 160
 QY 534 TGATCAACGAGACCTTTTGGCCGCTTTTGGCTCAGCTCCCTCAGGCGAGGCGG 593
 DB 161 TrrplethrgInCysPheLeuProValPheLeuAlaGlnProProSerGlyGlnArgArg 180

RESULT 10

AAB69946
 ID AAB69946 standard; Protein; 180 AA.

AC AAB69946;

XX 27-APR-2001 (first entry)

XX Human NY-ESO-1 protein.

XX Human; NY-ESO-1; HLA; human leukocyte antigen; CTL; cytotoxic T cell;

KM HLA-A2; HLA-DR5; melanoma; adenocarcinoma; bladder carcinoma;

KM non-small cell lung carcinoma; tumour status determination.

OS Homo sapiens.

XX WO200107917-A1.

XX 01-FEB-2001.

XX 14-JUL-2000; 2000WO-US19220.

XX 23-JUL-1999; 99US-0359503.

XX (LUDM-) LUDMIG INST CANCER RES.

PA (SLOK) SLOAN KETTERING INST CANCER RES.

PA (CORR) CORNELL RES FOUND INC.

XX Jager E, Stockert E, Old LJ, Knuth A, Chen Y, Scanlan M;

XX WPI; 2001-182822/18.

XX N-PSDB; AAF58634.

XX Example 5; Fig 3; 50bp; English.

XX The present sequence is human NY-ESO-1 protein. It is provided in a

CC specification relating to a method for determining the status of a

CC cancerous condition in a patient with a tumour that expresses NY-ESO-1.

CC The method comprises assaying a sample taken from the patient for

CC antibodies that specifically bind to the NY-ESO-1 and comparing the

CC value obtained to a prior value obtained from assay of a prior sample

CC taken from the patient. Any difference between the values is indicative

CC of a change in status of the cancerous condition. The method is useful

CC for determining whether a cancerous condition is progressing, regressing

CC or remaining stable, in particular in patients receiving treatment for a

CC melanoma, adenocarcinoma, non-small cell lung carcinoma or bladder

CC carcinoma.

XX Sequence 180 AA;

Alignment Scores:

Pred. No.: 3.71e-66 length: 180
 Score: 959.00 Matches: 180
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 68.35% Indels: 0
 DB: 22 Gaps: 0

QY 54 ATGACAGCCGAGAGCCGCGGAGCAGAGGAGGTTGACGCGGATGCTGATGCGCCAGAGGC 113
 DB 1 MetGlnAlaGlnGlyArgGlyThrGlyGlySerThrGlyAspAlaAspGlyProGlyGly 20
 QY 114 CCGGATTCCTGATATGCGCCAGGGGAGCATGCTGCGGCGCCAGAGAGGCGGCGCAGC 173
 DB 21 ProGlyIleProAspGlyProGlyGlyAsnAlaGlyIleProGlyGlyAlaGlyAlaThr 40
 QY 174 GCGCGCAGAGGTCCCGCGGCGCAGGCGCAGCAAGGCGCTGCGGCGGAGAGGCGCC 233
 DB 41 GlyIylArgGlyProArgGlyAlaGlyAlaAlaArgAlaSerGlyProGlyGlyAla 60
 QY 234 CCGCGGCGGTCGCGCATGCGGCGGCGGCTTCAGGCGTGAATGATGCTGAGATGCGGCGCC 293
 DB 61 ProArgGlyProHisGlyAlaAlaAlaSerGlyLeuAsnGlyCysArgCysGlyAla 80
 QY 294 AGGCGCGCGAGAGCGCGCTGCTTGAATTCACCTGCGCATGCTTTCGCGCAGCCCATG 353
 DB 81 ArgGlyProGlnSerArgLeuLeuGlnPheThrLeuAlaMetProPheAlaThrProMet 100
 QY 354 GAGCAGAGCTGCGCGCAGAGAGCTGCGCAGAGATGCGCCACCGCTTCCGTCGAGG 413
 DB 101 GluAlaGlnLeuAlaArgArgSerLeuAlaGlnAspAlaProProLeuProValProGly 120
 QY 414 GTGCTTCTGAAGAGATTCACTGTGTCCGCGCAACTGACTATCCGAGTCACTGCTGCA 473
 DB 121 ValLeuLeuGlyGlnPheThrValSerGlyAsnIleLeuThrIleArgLeuThrAlaAla 140
 QY 474 GACCAACCGCAACTGCAGCTCTCCATCAGCTCTGTCTCCAGAGCTTTCCCTGTATG 533
 DB 141 AspHisArgGlnLeuGlnLeuSerIleSerSerCysLeuGlnLeuSerIleuMet 160
 QY 534 TGATCAACGAGACCTTTTGGCCGCTTTTGGCTCAGCTCCCTCAGGCGAGGCGG 593
 DB 161 TrrplethrgInCysPheLeuProValPheLeuAlaGlnProProSerGlyGlnArgArg 180

RESULT 11

AAB84818
 ID AAB84818 standard; Protein; 180 AA.

XX AAB84818;

XX 08-MAY-2002 (first entry)

XX Human NYNSOLA consensus sequence.

XX Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;

KM viral infection; human immunodeficiency virus; melanoma;

KM bacterial infection; Salmonella; Legionella; parasitic infection;

KM Trypanosoma; Toxoplasma; Giardia.

OS Homo sapiens.

XX WO200190197-A1.

XX 29-NOV-2001.

XX 25-MAY-2001; 2001WO-AU00622.

XX 26-MAY-2000; 2000AU-000761.

XX (AUSU) UNIV AUSTRALIAN NAT.

XX Thomson SA, Ramshaw JA;

XX WPI; 2002-147575/19.

XX New synthetic polypeptides having several different segments of at

PT least one parent polypeptide linked together differently compared to

PT the linkage in the parent polypeptide, for inducing immune response

PT against a pathogen or cancer

Db	1	MeGlnAlaGluGlyYArgGlyThrhGlyGlySerThrGlyAlaSerGlyProGlyGly	20		
QY	114	CCGCGCATTCCTGATATGCCCCAGGGGGCAATGCTGGCCGGCCCAAGAGAGGGGGTGGCAAG	173		
Db	21	ProGlyIYIleProAlaSerGlyProGlyGlyYAsnAlaGlyGlyProGlyGlyAlaGlyAlaThr	40		
QY	174	GGCGGAGAGAGTCCCGGGGCGCAGGGGAGCGAAGGGCCCTGGGGCCGGGAGAGAGCGCC	233		
Db	41	GlyGlyYArgGlyYProArgGlyYAlaGlyAlaAlaSerGlyYleuAmGlyCySerGlyArgGlyYAla	60		
QY	234	CCGCGGGGCTCCCGATGCGCGCGCGCTTCAGGGCTGAATGGATGCTGCAGATCGGGGGCC	293		
Db	61	ProArgGlyYProHISelYalYAlaAlaSerGlyYleuAmGlyCySerGlyArgGlyYAla	80		
QY	294	AGGGGGGCGGAGAGCCCGCTGCTTGATGTTCTAACCCTCGGCATGCGCTTTGGGACACCCGATG	353		
Db	81	ArgGlyYProGlyIleSerArgIleuGluIleuGluPheYrYleuAlaIleMetProPheAlaIleThrProMet	100		
QY	354	GAGCAGAGAGCTGGCCCGGAGAGAGCTGGCCGACGATGGCCGACGCGCTTCCGGGACAGG	413		
Db	101	GluAlaGluIleuAlaArgArgSerIleuAlaGluIleuAlaProProIleuProValProGly	120		
QY	414	GTGCTTCTGAGAGAGTTCACGTGTGTCGGGCAACATACTGACTATCGAGTGACTGTGCA	473		
Db	121	ValIleuIleuYSerGluPheThrValSerGlyAsnIleIleuThrIleArgIleuThrAlaIa	140		
QY	474	GACCAACGCCCAACTGGAGAGCTTCACATGAGCTCTGTCCGACGAGCTTCCGTTGATG	533		
Db	141	AspIleArgGluIleuGluIleuSerIleSerCySerGluIleuIleuSerIleuIleuMet	160		
QY	534	TGGATCAACGAGTGTCTTGTCCCGGTGTTTGGCTCAGCCTCCCTCAGGGCAGAGGCGC	593		
Db	161	TrpIleThrGlnCySerIleuProValPheIleuAlaGlnProProSerGlyGlnArgArg	180		
RESULT 9					
AAU01535	AAU01535 standard, Protein, 180 AA.				
XX	AAU01535;				
ID	18-JUL-2001 (first entry)				
XX	18-JUL-2001 (first entry)				
DE	Human NY-ESO-1 tumour rejection antigen precursor protein.				
XX	NY-ESO-1; human; tumour rejection antigen precursor; SSX-2; MHC Class II;				
XX	major histocompatibility complex; helper T cell; HLA-DR; cancer;				
KW	human leukocyte antigen-determining region; disease progression;				
KW	disease regression; disease onset; body tissue; body fluid; enzyme label;				
KW	radioactive label; monoclonal antibody.				
XX	Homo sapiens.				
OS					
XX	Key	Location/Qualifiers			
FT	Modified-site	7 /note= "Myristoylated"			
FT	Modified-site	9 /note= "Myristoylated"			
FT	Modified-site	11 /note= "Myristoylated"			
FT	Modified-site	98 /note= "Phosphorylated"			
FT	Modified-site	134 /note= "Phosphorylated"			
FT	Modified-site	138 /note= "Phosphorylated"			
FT	Modified-site	/note= "Phosphorylated"			
XX	MO200123560-A2.				
PD	05-APR-2001.				
XX	26-SEP-2000; 2000OWO-US26411.				
PF					
XX					

XX	29-SRP-1993; 99US-0408036.
XX	(LUDW-) LUDWIG INST CANCER RES.
XX	
PI	Tureci O, Sahin U, Pfeunderschuh M;
DR	WFI, 2001-266156/27.
DR	N-PsDB; AAS02254.
XX	
PT	Polyepitides binding to major histocompatibility complex class II human
PT	leukocyte antigen-determining region molecule having amino acid
PT	sequence found in tumour rejection antigen precursor used for
PT	stimulating proliferation of helper T cells -
PS	Claim 4, Fig 3; 62pp; English.
CC	The sequence represents a human NY-ESO-1 tumour rejection antigen
CC	precursor. NY-ESO-1 and SSx-2 polyepitides, or fragments of, bind to
CC	major histocompatibility complex (MHC) Class II molecules such as human
CC	leukocyte antigen-determining region (HLA-DR) molecules and stimulate
CC	proliferation of helper T cells. The peptides can be administered to an
CC	HLA-DR positive subject in order to stimulate the helper T cells. An MHC
CC	Class II HLA-DR-NY-ESO-1/SSx-2 complex expressed on the surface of a cell
CC	or present in free form is useful for this stimulation. The nucleic acid
CC	is useful for screening for a cancerous condition, which involves
CC	contacting a subject sample to a cell line transfected with the
CC	immunoreactive cell (helper T cell), where interaction is indicative of
CC	cancer. In addition, a sample from a patient (for example, a body fluid
CC	or tissue) can be monitored for the amount of the complex present in the
CC	bloodstream. This is useful for determining regression, progression or
CC	onset of a cancerous condition. The method involves contacting the sample
CC	with a radioactive labelled or enzyme labelled monoclonal antibody which
CC	specifically binds with the complex.
SQ	Sequence 180 AA:
Alignment Scores:	
Pred. No.:	3,71e-66 Length: 180
Score:	959.00 Matches: 180
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	68.35% Indels: 0
DB:	22 Gaps: 0
US-10-023-182-1 (1-752) X AAU01535 (1-180)	
CY	54 ATGTCAGGCCGAAGGGCGGGGCACAGGGGGCTTCAAGGGCATGTGATGCGCCAGAGAGGC 113
Dd	1 MetGlnAlaGluGlyArgGlyThrIleYglSerThrGlySapAlaApGlyProGlyGly 20
CY	114 CCTGGCATTTCTGATGAGCGCCACAGGGGGCAATGCTGCGCGGCCCCAGAGAGGGCGGTCACG 173
Dd	21 ProGlyIleProAspGlyProGlyYglYasnAlaGlyIleProGlyGluAlaGlyAlaThr 40
CY	174 GGCGGAGAGAGTCCCCGGGGCGAGGGGCACAGAGGGCTCGGGGGCGGGAGAGAGGGCC 233
Dd	41 GlyIleArgGlyProArgGlyAlaGlyAlaAlaAlaArgAlaSerGlyProGlyGlyGlyAla 60
CY	234 CCCCGGGGATTCGGATGAGCGCGCGGCTTCAGAGGGTGATGATGATCTGCAGATGCGGGCC 293
Dd	61 ProArgGlyProHieGlyYglYAlaAlaSerGlyLeuSngIleCySArGySGlyAla 80
CY	294 AGGGGGCCGGAGAGCCGCTGCTTAAGTTAACTGCTGCGCATGCTTTOGGAGCACCCATG 353
Dd	81 ArgGlyIleProGlnSerTrgLeuLeuGlnHeTherYleuAlaMetProPheAlaIlePromet 100
CY	354 GAAGCAGAGAGTGGCCCGCAGAGAGCTGGCCAGAGATGGCCACCGCTTCCGATGCCAGG 413
Dd	101 GluAlaGluLeuAlaIlaArgTrgSerIleuAlaGlnSapAlaPropoleuProValProGly 120
CY	414 GTGCTTCTGAAGAATTCACTGTGTCCGGCAACTATGACTATACCACTGACTGCTGCA 473
Dd	121 ValIeuLeuYgluInHeTherValSerIleYasnIleuThrIleArgLeuThrAlaAla 140

PR 22-FEB-2000; 2000US-0510635.
XX (LUDW-) LUDWIG INST CANCER RES.
XX
PI Lethe B, Boon-Falleur T;
XX
XX WPI; 2001-550091/61.
DR N-PSDB; AAH75118.
XX
PT Genomic sequences of tumour associated antigen EY-ESO-1 (LAGE-2) useful
for diagnosing testicular tumours -
XX
XX Example 5; Fig 3; 50pp; English.
XX
XX The present sequence represents cancer testis tumour antigen NY-ESO-1
CC (also called LAGE-2). NY-ESO-1 is a molecule that is processed to at
CC least one human leukocyte antigen (HLA) binding peptide, which binds
CC to Class I and Class II major histocompatibility complex (MHC).
CC NY-ESO-1 is expressed in tumour mRNA and in testis, but not normal
CC colon, kidney, liver or brain tissue. The presence or level of expression
CC of NY-ESO-1 may be assayed for the diagnosis of cancer, especially
CC testis tumours.
XX
SQ Sequence 180 AA;

Alignment Scores:
Pred. No.: 3.71e-66 Length: 180
Score: 959.00 Matches: 180
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.35% Indels: 0
DB: 22 Gaps: 0

US-10-023-182-1 (1-752) x AAG67164 (1-180)

QY 54 ATGCAGGCCGGAAGGCGCGGGGACACAGGGGATTGCGAGCGGATGCTGATGCGCCAGAGAGGC 113
DB 1 MetGlnIaGlnGlyArgGlyThrGlyGlySerThrGlyAspAlaAspGlyProGlyGly 20
QY 114 CTTGGCATTCCTGATGCGCGCGGCGCAATGCTGGCGGCCAGAGAGGCGGGTCCACG 173
DB 21 ProGlyIleProAspGlyProGlyGlyAsnAlaGlyGlyProGlyGlnAlaGlyAlaThr 40
QY 174 GCGCGAGAGGTTCCCGCGGGCGACAGGGGACAGAGGCTCTCGGGCGCGGAGAGAGGCGC 233
DB 41 GlyIlyArgGlyProArgGlyAlaGlyAlaAlaArgAlaSerGlyProGlyGlyGlyAla 60
QY 234 CCGCGGGGTCGCGATGCGCGCGCGGCTTCAGGGTGAATGATGCTGCAATGCGGGGCC 293
DB 61 ProArgGlyProHISGlyGlyAlaAlaSerGlyLeuAsnGlyCysArgGlyAla 80
QY 294 AGGGGGCGCGAGAGCGCGCTGCTGAGTTCTACCTCGCCATGCTTCGCGACACCATG 353
DB 81 ArgGlyProGlnSerAlaGlyLeuGlnPheGlyIleuAlaMetProPheAlaThrProMet 100
QY 354 GAAGCAGAGCTGCGCGCGAGAGGCTTGCGCCAGAGATCCCAACCGCTTCCGTCGACAGG 413
DB 101 GlnAlaGlnLeuAlaArgArgSerLeuAlaGlnAspAlaProProLeuProValProGly 120
QY 414 GTGCTTCTGAAGAGGTTCACTGTTCCGGCAACTACTGACTGATCCGACTGACTGCA 473
DB 121 ValIleuLeuYsgIlePheThrValSerGlyAsnIleuThrIleArgLeuThrAlaAla 140
QY 474 GACCAACCGCCACTGCAAGCTCTCCATAGCTCTGTTCCAGAGCTTTCCTGTGATG 533
DB 141 AspHisArgGlnLeuGlnLeuSerIleSerSerGlyLeuGlnGlnLeuSerIleuMet 160
QY 534 TGATGACGAGAGTCTTTCGCGCGGCTTTTGGCTCAGAGCTCCCTCAGGGGACAGAGCGC 593
DB 161 TrpIleThrGlnCysPheLeuProValPheLeuAlaGlnProProSerGlyGlnArgArg 180

RESULT 8
AAE07714

ID AAE07714 standard; Protein; 180 AA.
XX
AC AAE07714;
XX
DT 06-NOV-2001 (first entry)
XX
XX Human NY ESO-1 protein.
DE
XX Human; major histocompatibility complex; MHC; vaccine; metastasis;
KW class II restricted T cell epitope; MHC-II epitope; cancer antigen;
KW NY ESO-1 protein; CD4+ T lymphocyte; human leucocyte antigen; HLA;
KW tumour-specific humoral-mediated immunity; cancer; cytotoxic;
KW immunotherapy.
XX
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT Misc-difference 45..47
FT /note= "Encoded by CCGGGGCGC"
PN WO200155393-A2.
XX
XX 02-AUG-2001.
PD
XX 26-JAN-2001; 2001WO-US02765.
PP
XX 28-JAN-2000; 2000US-0179004.
PR 29-SEP-2000; 2000US-0237107.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Wang R, Rosenberg SA, Zeng G;
XX WPI; 2001-496951/54.
XX N-PSDB; AAD14179, AAD14180.
DR
XX
PT New NY-ESO cancer peptide or MHC class II restricted T cell epitopes,
PT useful as immunogen and vaccine for inhibiting cancer in a mammal or as
PT protection from metastasis -
XX
XX Example 1; Fig 1; 134pp; English.
XX
XX The invention relates to the identification and isolation of major
CC histocompatibility (MHC) class II restricted T cell epitope (MHC-II
CC epitope) derived from the cancer antigen, NY ESO-1. The MHC-II epitopes
CC from NY ESO-1 are recognised by CD4+ T lymphocytes in an human leucocyte
CC antigen (HLA) class II restricted manner, in particular HLA-DR or HLA-DP
CC restricted. The products of the gene are promising candidates for
CC immunotherapeutic strategies for the prevention, treatment and diagnosis
CC of patients with cancer. The cancer epitopes are useful as immunogen and
CC vaccine to inhibit or to prevent cancer in a mammal by eliciting CD4+ T
CC lymphocytes resulting in protection of the recipient from development of
CC cancer and protection from metastasis, or by inhibiting the growth of
CC cells expressing the NY-ESO-1 gene product. The cancer peptides are also
CC useful as diagnostic agent to detect the presence of cancer, to enhance
CC the generation of antibody and/or CD8+ T cell responses against any
CC given target antigen and/or hapten and to induce tumour-specific
CC humoral-mediated immunity against cancer. The present sequence is
CC human NY ESO-1 protein.
XX
SQ Sequence 180 AA;

Alignment Scores:
Pred. No.: 3.71e-66 Length: 180
Score: 959.00 Matches: 180
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.35% Indels: 0
DB: 22 Gaps: 0

US-10-023-182-1 (1-752) x AAE07714 (1-180)

QY 54 ATGCAGGCCGGAAGGCGCGGGGACACAGGGGATTGCGAGCGGATGCTGATGCGCCAGAGAGGC 113

FT Peptide /note= "Peptide presented by MHC Class I HLA-B7"
 FT 113..112
 FT /note= "Peptide presented by MHC Class I HLA-B7 and
 FT HLA-B52"
 FT Peptide 115..124
 FT /note= "Peptide presented by MHC Class I HLA-A3"
 FT 118..126
 FT /note= "Peptide presented by MHC Class I HLA-B35"
 FT 124..133
 FT /note= "Peptide presented by MHC Class I HLA-B52"
 FT 125..133
 FT /note= "Peptide presented by MHC Class I HLA-A24"
 FT 138..147
 FT /note= "Peptide presented by MHC Class I HLA-B8"
 FT 139..147
 FT /note= "Peptide presented by MHC Class I HLA-B7"
 FT 145..153
 FT /note= "Peptide presented by MHC Class I HLA-A24 and
 FT HLA-B52"
 FT Peptide 153..162
 FT /note= "Peptide presented by MHC Class I HLA-B52"
 FT 154..162
 FT /note= "Peptide presented by MHC Class I HLA-B52"
 FT 154..163
 FT /note= "Peptide presented by MHC Class I HLA-B52"
 FT 156..167
 FT /note= "Peptide (AAV52434) presented by MHC Class I
 FT HLA-A2"
 FT Peptide 158..166
 FT /note= "Peptide presented by MHC Class I HLA-A3"
 FT 159..167
 FT /note= "Peptide presented by MHC Class I HLA-A3"
 FT 162..170
 FT /note= "Peptide presented by MHC Class I HLA-B52"
 FT Peptide
 XX WO953938-A1.
 XX 28-OCT-1999.
 XX 24-MAR-1999; 99WO-US06875.
 XX PF 17-APR-1998; 98US-0062422.
 XX PR 02-OCT-1998; 98US-0165546.
 XX (LUDM-) LUDWIG INST CANCER RES.
 XX PA Stockert E, Jager E, Chen Y, Scanlan M, Alexander K, Old LJ;
 PI Gure A, Rittner G;
 XX WPI: 2000-038483/03.
 XX N-ESDB; AA238380.
 XX Novel peptides which bind to MHC class I and MHC class II molecules,
 PT useful for therapeutic and diagnostic purposes -
 PT Claim 30; Fig 3; 49pp; English.
 XX This sequence represents a human tumour antigen, NY-ESO-1, the cDNA
 CC encoding which was isolated from an oesophagus squamous cell cancer cDNA
 CC library. Tissue localisation studies revealed it to be expressed at
 CC high levels in normal ovary and testis but not in normal colon, kidney,
 CC liver, brain, oesophagus and skin. It was expressed in certain tumours
 CC and tumour cell lines with some degree of frequency - these included
 CC melanoma specimens and cell lines, and breast and bladder cancer
 CC specimens, with expression in other tumour types being sporadic.
 CC Peptides derived from NY-ESO-1 are bound by both MHC (major
 CC histocompatibility complex) Class I and Class II molecules for
 CC presentation to T-cells. Peptides AAV52431-Y52434 bind to Class I HLA-A2
 CC molecules, thereby stimulating proliferation of cytotoxic T-cells, while
 CC peptides AAV52435-Y52440 bind to class II HLA-DR53 molecules, stimulating
 CC helper T-cell proliferation. The peptides derived from NY-ESO-1 may be
 CC used in methods and compositions used for the treatment, diagnosis and
 CC prevention of cancers (such as melanoma, breast cancer, prostate cancer,

CC lung cancer, hepatoma, ovarian cancer, thyroid cancer, bladder cancer,
 CC or lymphoma) and to stimulate the proliferation of T cells.
 XX SQ Sequence 180 AA;
 Alignment Scores:
 Pred. No.: 3,716-66 Length: 180
 Score: 959.00 Matches: 180
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 68.35% Indels: 0
 DB: 21 Gaps: 0
 US-10-023-182-1 (1-752) x AAV52430 (1-180)
 QY 54 ATGCAGGCCGAGAGCCGCGGCGACAGAGGGCTTCGACGGGCGATGCTGATGCCCGAGAGGC 113
 DB 1 MetGlnAlaGlnGlyValArgGlyThrGlyGlySerTrpGlyAspAlaAspGlyProGlyGly 20
 QY 114 CCGGATTCCTGATGCGCCAGAGGGCGCATGCTGCGCGCCAGAGAGCGCGTGCACG 173
 DB 21 ProGlyIleProAspGlyProGlyGlyAsnAlaGlyGlyProGlyGlyAlaThr 40
 QY 174 GCGCGCAGAGGTCCCGCGGCGCCAGAGGGCGACAGAGGGCTCGGGGCCCGAGAGAGCGCC 233
 DB 41 GlyGlyArgGlyProArgGlyValaGlyAlaAlaArgAlaSerGlyProGlyGlyVala 60
 QY 234 CCGCGGGGTCGCGATGCGCGCGCGGCTTCAGGGCTGATGATGATGATGCGAGATGCGGGGCC 293
 DB 61 ProArgGlyProHisGlyValaAlaSerGlyLeuAsnGlyCysCysArgGlyVala 80
 QY 294 AGGAGGCGCGAGAGCGCGCTGCTTGAATCTACCTTCGCGATGCTTTGCGCACCCGATG 353
 DB 81 ArgGlyProGlyIleSerArgLeuGlnPheTrpAlaMetProPheAlaThrProMet 100
 QY 354 GAAGCAGAGCTGGCCCGCGAGAGCGCTGCGCGCGAGATGCCCGCTTCGCGAGAGG 413
 DB 101 GlnAlaGlnLeuAlaArgArgSerLeuAlaGlnAspAlaProProLeuProValProGly 120
 QY 414 GTGCTTCGAGAGGATTCAGTGTGTCGCGCAACATGACTGATTCGACTGCTGCA 473
 DB 121 ValLeuLeuGlySerValSerValSerGlyValSerGlyValSerGlyValSerGlyVal 140
 QY 474 GACCAACCGCCACACTGAGAGCTTCACAGCTCCTGCTCCAGAGTTCCGTTGATG 533
 DB 141 AspHisArgGlnLeuGlnLeuSerIleSerCysLeuGlnGlnLeuSerLeuLeu 160
 QY 534 TGGATCAGCAGAGTCTTCTGCGCGGCTTTGGCTCAAGCTTCCTCAGAGGAGAGCGCG 593
 DB 161 TrpIleThrGlnGlySheLeuProValPheLeuAlaGlnProProSerGlyGlnArgGly 180
 RESULT 7
 AAG67164
 ID AAG67164 standard; Protein; 180 AA.
 XX AAG67164;
 XX DT 13-NOV-2001 (first entry)
 XX Amino acid sequence of cancer testis tumour antigen NY-ESO-1.
 XX DE Cancer testis tumour antigen; NY-ESO-1; LAGE-2; human leukocyte antigen;
 XX KW HLA; HLA binding peptide; major histocompatibility complex; MHC;
 XX tumour; cancer; testis tumour.
 XX OS Homo sapiens.
 XX MO200162917-A1.
 XX 30-AUG-2001.
 XX 22-JAN-2001; 2001WO-US02126.


```

Db 101 GlnuagluuAlaArgArseruulaglnuAlaProPheuProValProGly 120
Qy 414 GTGCTTGAAGAGATTGCTGTGTCGCGCAACTACTGATTCGACTGCTGCA 473
Db 121 ValLeuLeuYsgluPheThrValSerGlyAsnIleLeuThrIleArgLeuThrAlaAla 140
Qy 474 GACCAACCCCACTGCACTCTTCATGCTCCGCTCCGACGAGCTTCCCTGTGATG 533
Db 141 AsphIsArgInLeuGlnLeuSerIleSerCysLeuGlnIleuSerLeuLeuMet 160
Qy 534 TGGATTCACGAGTGTCTTCTGCGCTGTTTGGCTCAGCCTCCCTCAGGCGACAGGCGC 593
Db 161 TrpIleThrGlnCysPheLeuProValPheLeuAlaGlnProProSerGlyGlnArgArg 180

RESULT 4
AAB03154
ID AAB03154 standard; Protein; 180 AA.
AC AAB03154;
XX 23-OCT-2000 (first entry)
DT
XX Human oesophageal cancer-associated antigen NY-ESO-1.
XX
XX Oesophageal cancer associated antigen; NY-ESO-1; human;
XX immunogen; oesophageal carcinoma; melanoma; ovary; testis;
XX transmembrane domain; antibody; diagnostic marker; drug delivery target.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 7 /note= "Potential N-myristoylation site"
XX Modified-site 9 /note= "Potential N-myristoylation site"
XX Modified-site 11 /note= "Potential O-phosphorylation site"
XX Modified-site 98 /note= "Potential O-phosphorylation site"
XX Modified-site 134 /note= "Potential O-phosphorylation site"
XX Modified-site 138 /note= "Potential O-phosphorylation site"
XX Modified-site 152..172 /note= "Potential transmembrane domain"
XX Domain
XX
XX US6069233-A.
XX
XX 30-MAY-2000.
XX
XX 26-JAN-1998; 98US-0013150.
XX
XX 03-OCT-1996; 96US-0725381.
XX
XX (SLOK ) SLOAN KETTERING INST CANCER RES.
XX (CORR ) CORNELL RES FOUND INC.
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Scanlan M, Gure AO, Chen Y, Tureci O, Sahin U, Pfreundschuh M;
XX Old LJ;
XX
XX MPI; 2000-410880/35.
XX N-PSDB; AAA61483.
XX
XX New isolated oesophageal cancer-associated antigen useful as markers for
XX producing antibodies and as targets for identifying abnormal
XX conditions, e.g. infections and cancer -
XX
XX Example 5; Fig 3; 9pp; English.
XX
XX This sequence represents a human oesophageal cancer-associated antigen,
XX NY-ESO-1. The CDNA encoding this sequence was isolated from a

```

```

CC CCNA library prepared from a specimen of well-to-moderately
CC differentiated squamous-cell cancer of the oesophagus. Expression
CC analysis demonstrated that NY-ESO-1 was expressed in oesophageal
CC carcinoma, certain melanoma cell lines and in normal ovary and testis
CC tissue, but not in normal colon, kidney, liver or brain tissue. Analysis
CC of the amino acid sequence of the protein indicates that the protein has
CC a transmembrane domain, several N-myristoylation sites and
CC O-phosphorylation sites and that it contains antigenic sequences in the
CC N-terminal half of the protein. The antigen is useful as an immunogen
CC when combined with an adjuvant, in both precursor and post-
CC translationally modified forms, and may be used to generate anti-NY-ESO-1
CC antibodies. It can also be used as a diagnostic marker for oesophageal
CC cancer, and can be utilized as a marker for the targeted delivery of
CC therapeutic agents to oesophageal cancer cells. It can also be used to
CC generate diagnostic or therapeutic agents.
XX
XX Sequence 180 AA:
XX
XX Alignment Scores:
XX Pred. No.: 3.71e-66 Length: 180
XX Score: 959.00 Matches: 180
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 68.35% Indels: 0
XX DB: 21 Gaps: 0

US-10-023-182-1 (1-752) x AAB03154 (1-180)
Qy 54 ATGCAGCCGGAAGGCGCGGCGACAGGCGGTTGACGCGGATGCTGATGCGCCAGAGGC 113
Db 1 MetGlnAlaGlnGlyArgGlyThrGlySerThrGlyAspAlaAspGlyProGlyGly 20
Qy 114 CCGGATTCCTGATGAGCCCGGCGGCAATGCTGCGCGCCCGGAGAGCGGCGCC 173
Db 21 ProGlyIleProAspGlyProGlyGlyAsnAlaGlyGlyProGlyGlnAlaGlyAlaThr 40
Qy 174 GCGCGAGAGGTCGCCGCGGCGCGGCGGCGAGGAGGAGGCGGCGCGGCGCGGCG 233
Db 41 GlyIleArgGlyProArgGlyAlaGlyAlaAlaAlaAlaAlaSerGlyProGlyGlyAla 60
Qy 234 CCGCGGCGGTCGCCGATGCGCGCGCGGCTTCAAGGCTGATGATGCTGCAAGATGCGGCGC 293
Db 61 ProArgGlyProIleGlyAlaAlaAlaSerGlyLeuAlaGlyCysArgGlyAla 80
Qy 294 AGGCGGCGCGGAGAGCGCGCTGCTGATGCTGCTGATGCTGCTGCGGATGCTGCGGACCCATG 353
Db 81 ArgGlyProGlnSerArgLeuLeuGlnPheThrIleAlaMetProPheAlaThrProMet 100
Qy 354 GAAGCAGAGCTGCGCGCGGCGGAGGCTGCGGCGGATGCGGCGGCTTCCGTCAGG 413
Db 101 GlnuagluuAlaArgArseruulaglnuAlaProPheuProValProGly 120
Qy 414 GTGCTTGAAGAGATTGCTGTGTCGCGCAACTACTGATTCGACTGCTGCA 473
Db 121 ValLeuLeuYsgluPheThrValSerGlyAsnIleLeuThrIleArgLeuThrAlaAla 140
Qy 474 GACCAACCCCACTGCACTCTTCATGCTCCGCTCCGACGAGCTTCCCTGTGATG 533
Db 141 AsphIsArgInLeuGlnLeuSerIleSerCysLeuGlnIleuSerLeuLeuMet 160
Qy 534 TGGATTCACGAGTGTCTTCTGCGCTGTTTGGCTCAGCCTCCCTCAGGCGACAGGCGC 593
Db 161 TrpIleThrGlnCysPheLeuProValPheLeuAlaGlnProProSerGlyGlnArgArg 180

RESULT 5
AAY70862
ID AAY70862 standard; Protein; 180 AA.
AC AAY70862;
XX 31-JUL-2000 (first entry)
XX
XX Human tumour antigen, NY-ESO-1 protein.
XX

```

CC condition are also described. The NY-ESO-1 protein and peptides derived
CC from it can be used for diagnosis and treatment of cancers and to monitor
CC the efficacy of a therapeutic regime.

XX Sequence 180 AA:

Alignment Scores:

Score:	3,71e-66	Length:	180
Percent Similarity:	959.00	Matches:	180
Best Local Similarity:	100.00%	Conservative:	0
Query Match:	100.00%	Mismatches:	0
DB:	68.35%	Indels:	0
	19	Gaps:	0

US-10-023-182-1 (1-752) x AAW62584 (1-180)

```
QY 54 ATGCGAGCCGGAAGCCGCGGCGCAGAGGGGTTGACAGGGCGATGCTGATGCCCGAGAGGC 113
DB 1 MetGlnAlaGlnGlyArgGlyThrGlyGlySerThrGlyAspAlaAspGlyProGlyGly 20
QY 114 CTTGCATTCCTGATGCGCCAGAGGGGCAATGCTGCGCGCCAGAGAGGGCGGTGCCAG 173
DB 21 ProGlyIleProAspGlyProGlyGlyAsnAlaGlyGlyProGlyGlnAlaGlyAlaThr 40
QY 174 GCGCGCAGAGGTCCCGCGGGCGCAGAGGGGCTCGGGGCCCGGAGAGAGGCC 233
DB 41 GlyGlyArgGlyProArgGlyAlaGlyAlaAlaAlaSerGlyProGlyGlyAla 60
QY 234 CGCGGGGTCGCGATGCGCGCGCGGCTTCAGGGCTGATGATGATGCGAGATGCGGGGCC 293
DB 61 ProArgGlyProHisGlyGlyAlaAlaSerGlyLeuAsnGlyCysCysArgCysGlyAla 80
QY 294 AGGGGGCCGAGAGCGCGCTGCTTGAATCTTACTTCGCGCATGCTTTGGCGACCCATG 353
DB 81 ArgGlyProGlySerArgLeuGlnGlyPheThrIleAlaIleProPheAlaThrProMet 100
QY 354 GAAGCAGAGCTGCGCGCGAGAGCGCTGCGCGCGAGATGCCCGCTCCGTCGACAGG 413
DB 101 GlnAlaGlnLeuAlaArgArgSerLeuAlaGlnAspAlaProPheLeuProValProGly 120
QY 414 GTGCTTCTGAGAGATTCATCTGTGTCGCGCAACATACCTACTATCCGACTGCTGCA 473
DB 121 ValLeuLeuGlyGlnPheThrValSerGlyAsnIleLeuThrIleArgLeuThrAlaAla 140
QY 474 GACCAACCGGCACTGAGAGCTTCACATCAAGCTCCTGCTCCAGACGTTCCCTGTTGATG 533
DB 141 AspHisArgGlnLeuGlnLeuSerIleSerSerCysLeuGlnGlnLeuSerLeuLeuMet 160
QY 534 TGGATCAGCCAGTCTTCTTGCCTGTTGCTGCTCAAGCTCCCTCAAGGAGAGAGGCGC 593
DB 161 TrpIleThrGlnCysPheLeuProValPheLeuAlaGlnProProSerGlyGlnArgArg 180
```

RESULT 3

AA05965

ID AAY05965 standard; Protein; 180 AA.

XX AAY05965;

DT 16-AUG-1999 (first entry)

XX Human cancer antigen NY ESO-1/CAG-3 ORF1 protein.

XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer antigen; human; leukaemia;

XX non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;

XX metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;

XX uterine cancer; breast cancer; prostate cancer; ovarian cancer;

XX cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;

XX liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;

XX vaccine; ORF1.

XX Homo sapiens.

XX OS

XX W0918206-AA2.

```
XX 15-APR-1999.
PD 21-SEP-1998; 98WO-US19609.
XX 08-OCT-1997; 97US-0061428.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA Rosenberg SA, Wang RF;
PI WPI; 1999-277270/23.
XX N-PSDB; AAK58599.
XX Cancer antigen NY ESO1/CAG-3
PT Claim 4; Fig 3A; 88pp; English.
PS The present sequence represents the ORF1 protein encoded by
XX open reading frame 1 of the human ESO-1/CAG-3 (or CAG-3) gene.
XX CAG-3 is a new and potent tumour antigen capable of eliciting an
XX antigen specific immune response by T cells. Cancer peptides
XX comprising ORF1, ORF2 (see AAY05966), portions of these peptides and
XX their variants (see AAY05965-87), are useful as cancer vaccines that
XX protect the recipient from development of cancer. The invention
XX provides vectors and host cells (also useful as vaccines); a
XX method of diagnosis of cancer or precancer; a transgenic animal;
XX antisense oligonucleotides that inhibit expression of the cancer
XX peptide or tumour antigen; antibodies reacting with the CAG-3
XX cancer peptide, useful in diagnostic and detection assays; and
XX methods for preventing or inhibiting cancer by administering a
XX cancer peptide, with or without an H1A molecule. The cancer
XX peptides form part of, or are derived from, cancers such as primary
XX or metastatic melanoma, lymphoma, sarcoma, lung cancer,
XX liver cancer, leukaemia, uterine cancer, cervical cancer, bladder
XX cancer, kidney cancer and adenocarcinomas such as breast, prostate,
XX ovarian, pancreatic and thyroid cancers. Melanoma is treated by
XX inducing cancer-specific T cells in vitro for subsequent return to
XX a patient.
XX
XX Sequence 180 AA:
SQ Alignment Scores:
Pred. No.: 3,71e-66 Length: 180
Score: 959.00 Matches: 180
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.35% Indels: 0
DB: 20 Gaps: 0
US-10-023-182-1 (1-752) x AAY05965 (1-180)
QY 54 ATGAGGCCGGAAGCCGCGGCGCAGAGGGGTTGACAGGGCGATGCTGATGCCCGAGAGGC 113
DB 1 MetGlnAlaGlnGlyArgGlyThrGlyGlySerThrGlyAspAlaAspGlyProGlyGly 20
QY 114 CTTGCATTCCTGATGCGCCAGAGGGGCAATGCTGCGCGCCAGAGAGGGCGGTGCCAG 173
DB 21 ProGlyIleProAspGlyProGlyGlyAsnAlaGlyGlyProGlyGlnAlaGlyAlaThr 40
QY 174 GCGCGCAGAGGTCCCGCGGGCGCAGAGGGGCTCGGGGCCCGGAGAGAGGCC 233
DB 41 GlyGlyArgGlyProArgGlyAlaGlyAlaAlaAlaSerGlyProGlyGlyAla 60
QY 234 CGCGGGGTCGCGATGCGCGCGCGGCTTCAGGGCTGATGATGATGCGAGATGCGGGGCC 293
DB 61 ProArgGlyProHisGlyGlyAlaAlaSerGlyLeuAsnGlyCysCysArgCysGlyAla 80
QY 294 AGGGGGCCGAGAGCGCGCTGCTTGAATCTTACTTCGCGCATGCTTTGGCGACCCATG 353
DB 81 ArgGlyProGlySerArgLeuGlnGlyPheThrIleAlaIleProPheAlaThrProMet 100
QY 354 GAAGCAGAGCTGCGCGCGAGAGCGCTGCGCGCGAGATGCCCGCTCCGTCGACAGG 413
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_nzp model

Run on: February 10, 2004, 17:20:36 ; Search time 41.5 Seconds

(without alignments)
5752.404 Million cell updates/sec

Title: US-10-023-182-1

Perfect score: 1403
Sequence: 1 ATCTCGTGGGCGCTGACCT.....TAAACAGACTACGAAAA 752

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1107863 segs, 158726573 residues

Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+nzp,model -DEV=xlp
-O=/cgm2_1/USPRO.spool_p/US10023182/runat_10022004_171428_3849/asp_query.fasta_1.903
-DB=AGeneSeq_19Jun03 -QMT=fastran -SUFFIX=razg -MINMATCH=0.1 -LOOFCUT=0
-LOOEXT=0 -UNIT5=edit -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=Pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USE=US10023182 -CGN 1 1 81 @runat_10022004_171428_3849 -NCPU=6 -ICPU=3
-NO MAP -LARGOUTERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGCG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A.GeneSeq_19Jun03:*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	959	68.4	180	19	AAW69665
2	959	68.4	180	19	AAW62584
3	959	68.4	180	20	AAV05965
4	959	68.4	180	21	AAW03154
5	959	68.4	180	21	AAV70862
6	959	68.4	180	21	AAV52430
7	959	68.4	180	22	AAW67164
8	959	68.4	180	22	AAW07714
9	959	68.4	180	22	AAW01535
10	959	68.4	180	22	AAW69946
11	959	68.4	180	23	AAW84818
12	959	68.4	180	23	AAU11543
13	959	68.4	180	24	ABR48210
14	959	68.4	180	24	ABU56508
15	959	68.4	180	24	ABU56694
16	959	68.4	180	24	ABP74198
17	959	68.4	397	22	AAE13122
18	953	67.9	180	24	ABU64816
19	821	58.5	180	24	ABP74199
20	814	58.0	180	19	AAW69664
21	814	58.0	180	21	AAV70860
22	814	58.0	180	23	ABB78346
23	814	58.0	180	23	AAU84820
24	631.5	45.0	135	24	ABR48211
25	631.5	45.0	135	24	ABU56509
26	631.5	45.0	135	24	ABU56695
27	597	42.6	210	24	ABP74200
28	591	42.1	210	19	AAW69663
29	591	42.1	210	21	AAV70861
30	455	31.7	3541	23	AAU85130
31	443	30.9	109	21	AAV70854
32	285	20.3	58	20	AAV05966
33	285	20.3	58	21	AAV70863
34	285	20.3	58	23	AAU84819
35	203.5	14.6	330	19	AAW57645
36	203.5	14.6	408	17	AAW07539
37	203.5	14.6	408	22	AAW72737
38	203.5	14.6	408	22	AAW64007
39	202.5	14.5	561	14	AAW37739
40	202.5	14.5	561	17	AAW93249
41	202.5	14.5	561	19	AAW57650
42	202.5	14.5	777	14	AAW37740
43	202.5	14.5	777	17	AAW93250
44	202.5	14.5	777	19	AAW57651
45	201	14.4	357	17	AAW95115

ALIGNMENTS

RESULT 1	AAW69665
ID	AAW69665 standard; Protein; 180 AA.
AC	AAW69665;
XX	
DT	27-OCT-1998 (first entry)
XX	
DE	Human NY-ESO-1 protein sequence, formerly known as LL-1.2 clone.
XX	
KM	Human; LAGE-1; tumour associated protein; LL-1; diagnosis; tumour.
XX	
OS	Homo sapiens.
XX	
FN	MO9832855-A1.
XX	
PD	30-JUL-1998.
XX	

RA Shinkets L.J. to the EMBL/GenBank/DBJ databases.
 RU Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF494111; AAI56589.1; ATPase.
 DR InterPro: IPR003594; ATPase.
 DR InterPro: IPR002545; Chew.
 DR InterPro: IPR004105; H-kinase_dim.
 DR InterPro: IPR005467; His_kinase.
 DR InterPro: IPR002570; Hpt.
 DR Pfam: PF01584; Chew; 1.
 DR Pfam: PF02895; H-kinase_dim; 1.
 DR Pfam: PF02518; HATPase_c; 1.
 DR Pfam: PF01627; Hpt; 1.
 DR ProDom: PD003142; Hpt; 1.
 DR SMART: SM00260; Chew; 1.
 DR SMART: SM00387; HATPase_c; 1.
 DR SMART: SM00073; Hpt; 1.
 DR PROSITE: PSS0109; HIS_KIN; 1.
 SQ SEQUENCE 853 AA; 86305 MW; 386DC7218D67760C CRC64;

Alignment Scores:
 Pred. No.: 7.21e-06 Length: 853
 Score: 188.50 Matches: 78
 Percent Similarity: 33.21% Conservative: 10
 Best Local Similarity: 29.43% Mismatches: 80
 Query Match: 13.44% Indels: 97
 DB: 2 Gaps: 11

US-10-023-182-1 (1-752) x DBVQZ0 (1-853)

QY 36 GCGAGAGGCTCCGAGGATGCGAGAGCGGCGGATTCAGAGGCGAT 95
 DB 151 GIVGLYseraspGlyserGlyThrAlaGlyserGlyThrAlaGlyserGlyThr 170
 QY 96 GCTGATGCCCGAGAGGCGCTGGCAT 122
 DB 171 ThrGlyValaglyserGlyThrThrGlyGlyseraspGlyserGlyThrThr 190
 QY 123 -----CCTGATGCC-----CCA 134
 DB 191 GIVGLYValProaspGlyserGlyThrAlaGlyValAlaGlyserGlyThrThr 210
 QY 135 GCGGCGCATGCTCGCGCGCCGAGAGCGGCGGCGCGAGAGGTCCTCCCGGCGC 194
 DB 211 GIVGLYValaglyserGlyThrAlaGlyValaglyserGlyProThrGly 230
 QY 195 GCGAGGCGAGAGGCGCTCGGCGCGGAGAGCGCGCGGGGTCCGATGGCGGC 254
 DB 231 AlaGlyAlaSerProAlaSerThrAlaGlyValAla----- 243
 QY 255 GCGGCTTCAGGCGTGAATGATGCTGCGAGTGCAGGCGCGAGAGCGCGATG 314
 DB 244 AlaGlyserGlyThrArgSerCys-ValaspGlyserGly-----AlaAla 259
 QY 315 CTGAGTTCTACTCGCATGCTTTGGCAGC-----ACCATGAGAGAGAGCTG 365
 DB 259 aglyGlyserAlaSerProGlyValAla-SerGlyThrGlyValaspGlyAlaGlyProG 279
 QY 366 GCGCGAGAGAGCGCT-----GCGCGAGAGTGCACCGCTTCCCGTGGCCA 410
 DB 279 IYProAlaSerProAlaSerLeuMetasnGlyGlyProGlyThrSerAlaThrSerAlaP 299
 QY 411 GGGGTGCTTGAAGAAGTCACTGT----- 436
 DB 299 roasnGlyAlaAlaGlyThrSerAlaProLeuSerProaspGlyAlaThrGlyValY 319
 QY 437 -----GTCCGCGACATAGTACTATCCG---ACTGACTGCTGCAGACCAAC 479
 DB 319 alValProLeuSerThrArgArgAlaSeraspValProGlyThrAspAlaGlySerAla 339
 QY 480 CGCGACTGCAGCTCTGCATGAGCTCTGCTCCAGCAGCTTCCCTGTTGATGTGATC 539
 DB 339 IapProThrProaspLeuAlaasnSerLeuYsaIaAlaAla----- 352

QY 540 ACGAGTGTCTTGTGCCCCGTGTTTGGCTCAGCCTCCCTCAGAGGCGCTAAGCC 599
 DB 353 -----SerAlaProLeu-----ProGlyThrLeuAlaG 362
 QY 600 CAGCCTG-----CGCCCT 614
 DB 362 IArgTPrAlaValArgLeuArgIleAlaProThrCysGlnValProGlyValArgAlaP 382
 QY 615 TCCTAGGTGAT 625
 DB 382 heLeuValHis 385

Search completed: February 10, 2004, 17:36:14
 Job time : 49.5 secs

OX	NCBI_TaxId=29158;
RN	[1]
RP	SEQUENCE FROM N.A.
RE	MEDLINE=22038007; PubMed=12042339;
RA	Lucas J.M., Vaccaro E., Waite J.H.;
RT	A molecular, morphometric and mechanical comparison of the structural elements of byssus from Mytilus edulis and Mytilus galloprovincialis.";
RL	J. Exp. Biol. 205:1807-1817(2002).
DR	EMBL: AF48526; AAM34601.1;
DR	InterPro: IPR000087; Collagen.
DR	Pfam: PF01391; Collagen.7.
DR	Prodom; PD000007; Collagen; 2.
KM	Collagen.
SQ	SEQUENCE 922 AA; 80751 MW; CA281AE0975BBD4 CRC64;
Alignment Scores:	
Pred. No.:	5 08e-06 Length: 922
Score:	189.50 Matches: 63
Percent Similarity:	41.67% Conservative: 2
Best Local Similarity:	40.38% Mismatches: 68
Query Match:	13.51% Indels: 23
DB:	Gaps: 5 Gaps: 8
US-10-023-182-1 (1-752) x Q8WM53 (1-922)	
OY	33 CCGGGCAGAGGCTCCGGAGCCATGCAGGCCGAGGCGGACAGGGGGTTGCAGCGGC 92
Dd	163 ProGIymeProGIyGIyProGIyGIyProberGIyProGIythrGIyGIy---ProGIy 181
OY	93 GATGCTGATGGCCGAGGAGGCGCTTCCTGATGAGCCGAGGGGCAATGCT----- 146
Dd	182 GlmProAspGIyProGIyGIyProGIyGIyProGIyGIyProGIyGIyProGIymePro 201
OY	147 GGGGCGCCAGAGAAGGCGGGGTGCCACGGGCGGAGAGGCTCCC GGCGGCGAGGGCAGCA 206
Dd	202 GIyGIyProGIyGIyProGIyGIyProGIyGIyProGIyGIyProGIyGIyProGIy 221
OY	207 AGGGCTCGGGGCGGGAGAGAGCGCCCGGGGGTCCGATGGCGGCGGCTTCAGGG 266
Dd	222 AspGIyGIyIleProGIymeProGIyProAlaGIyProProGIyProAla----Gly 239
OY	267 CTGAATGATGCTGCAGATGCCGGGCGAGGGGCGGAGAGCCGCTTGACTTCATC 326
Dd	240 ProGIInGIy-----ProGIuGIyGluGlnGIyProAtgGIyAsg----- 252
OY	327 CTGGCCATGCTTTCGCCGACACCCTATGAGAGAGAGCTGGCGCCGACGAGGCTGGC---C 383
Dd	253 -----ThrProAlaGIyThrPro--GIyProProGIyAsnProGIyGluProGIyGlnA 270
OY	384 CAGGATGCCCACCGCTGCTCCGCTGCAGGGGTGCTTCGAAAGAGATTCACTGTGTCCGCG 443
Dd	270 IagIyAlaProGIyAlaProGIyAlaProGIy-----HisIaGIyLysH 285
OY	444 AACATACTGACTATCCGACT---GACTGCTGAGAACCCAGCCGCA 484
Dd	285 IagGIyThralaGIyProAlaGIyLysalagIyArGIyProGIyPro 299
RESULT 15	
ID	Q8VQZ0 PRELIMINARY; PRT; 853 AA.
AC	Q8VQZ0;
DT	01-MAR-2002 (TREMBLrel. 20, Created)
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE	Dife protein.
OS	Myxococcus xanthus.
OC	Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC	Cystobacteriales; Myxococcaceae; Myxococcus.
NC	NCBI_TaxId=34;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=DK1622;


```

Db      290 GlyProLeuProProGlyLProProProProValProGlyTYGlyLProProProGlyPro 309
Oy      109 CCTGGGCGCATGACATCGCCCTCGAACCCCTGTGCC----- 71
Db      310 ProProProGInGInGlyLProProProProProGlyLProPheProProArgProProGly 329
Oy      70 -----CGGCTTGGCGCTGATGAGTCCCGAGGCT---CTGCGCGGC 32
Db      330 ProLeuGlyProProLeuThreuThreuAlaProProProHisLeuProGly 345

RESULT 12
O16630
AC      Q16630      PRELIMINARY;      PRT;      551 AA.
DT      01-NOV-1996 (TREMBLrel. 01, Created)
DT      01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT      01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE      HPRBII-4 mRNA.
GN      HPRBII-4 OR HPRBII-7.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX      NCBI_Taxid=9606;
RN      (1)
RP      SEQUENCE FROM N.A.
RA      Fleischauer K.L.;
RL      Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
DR      EMBL; X67337; CAA47752.1; -
DR      EMBL; X67336; CAA47751.1; -
DR      Genew; HGNC:13871; CP5F6
DR      InterPro; IPR002965; P_Rich_extensn.
DR      InterPro; IPR005054; RNA_rec_moc.
DR      Pfam; PF0076; xrm; 1.
DR      PRINTS; PRO1217; PRICHEXTENSN.
DR      SMART; SMO0360; RRM; 1.
DR      PROSITE; PS50102; RRM; 1.
SQ      SEQUENCE 551 AA; 59208 MW; 407CB941B456A966 CRC64;

Alignment Scores:
Pred. No.:      5,656-06      Length:      551
Score:      190.00      Matches:      59
Percent Similarity:      39.87%      Conservative:      4
Best Local Similarity:      37.34%      Mismatches:      46
Query Match:      13.59%      Indels:      49
DB:      4      Gaps:      7

US-10-023-182-1 (1-752) x Q16630 (1-551)

Oy      412 CCTGGACGGGAAGCGGTGGGACATCTCGGCGCAGGCTCCTCGGCGCAGCTGTCTCC 353
Db      208 ProGlyAlaValProGlyGlyAspArgPhe-ProGlyProAlaGlyPro----- 223
Oy      352 ATGGGTGTGCGGAAGAGCATGCGAGGTAGAACTAAGACGGCGGTCTCGGCGCCCTG 293
Db      224 -----GlyGlyProPro-ProProP 230
Oy      292 GCCCGCATGTGACGATCATTCATCGCCTGAAAGCCGCGCG-----CATGCGACCC 239
Db      230 heProAlaGly-GInThrProProArgProProLeuGlyLProProGlyLProProGlyPro 249
Oy      238 CGGCGGGCGCTCTCCGCGC-----CCGAGGCGCTGTGTCGCCCTGCGCGCCGG 188
Db      250 ProGlyProProProProGlyGInValLeuProProProLeuAlaGlyProProAsnArg 265
Oy      187 GGA-----CCTCTG-----CGCCCGGTG 170
Db      270 GlyAspArgProProProProValLeuPheProGlyGInProPheGlyGInProLeu 285
Oy      169 GCAACCGCGCTCTCCTGGGCGCGCGACGATTGCCCCCTGGGCGCATGAGGAATCCAGGCGCT 110
Db      290 GlyProLeuProProProGlyLProProProProValProGlyTYGlyLProProProGlyLPro 309

```

QY		109	CCTGGGACCATAGACANMGCCGTGAACCCCTGTGCC-----	71
DY		310	ProProGInGLnglNglyProProProProProGlyProPhneProProArgProProGly	329
DB				
QY		70	-----CGGCCTTCGGCCCTCAGTGGCTCCGAGCCT---CTGCCCGGC	32
Db		330	ProLeuglyProProLeuthreuthraProProHisleuProGly	345
RESULT 13				
Q9F2NS	PRELIMINARY;	PRT:	889 AA.	
AC	Q9F2NS;			
DT	01-MAR-2001 (TrEMBLrel. 16, Created)			
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)			
DE	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)			
DS	Hypothetical protein SC03115.			
OS	SCO3115 OR SCF41.24C.			
OC	Streptomyces coelicolor.			
OC	Bacteria; Actinobacteriae; Actinobacteridae; Actinomycetales;			
OX	NCBI_TaxID=1902;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=A3(2);			
RA	Saunders D.C., Harris D.;			
RL	Submitted (SEP-2000) to the EMBL/Genbank/DDBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=A3(2);			
RA	Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;			
RL	Submitted (SEP-2000) to the EMBL/Genbank/DDBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=A3(2);			
RA	MEDLINE=97000351; PubMed=8643435;			
RL	Redenbach M., Kleiser H.M., Denapate D., Eichner A., Cullum J.,			
RA	Kinash H., Hopwood D.A.;			
RT	"A set of ordered cosmids and a detailed genetic and physical map for			
RL	the 8 Mb Streptomyces coelicolor A3(2) chromosome.";			
MoI	Microbiol. 21:77-96(1996).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=A3(2) / M145.			
RA	MEDLINE=21996410; PubMed=12000953;			
RL	Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,			
RA	Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleiser H.,			
RA	Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,			
RA	Croddin A., Festerman A., Gobie A., Hidalgo J., Hornsbay T., Howarth S.,			
RA	Huang C.-H., Kleiser T., Larke L., Murphy L., Oliver K., O'Neill S.,			
RA	Rabunowitsch E., Rajandream M.A., Rutherford K., Rutter S.,			
RA	Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,			
RA	Warren T., Metcortex A., Woodward J., Barrell B.G., Parkhill J.,			
RA	Hopwood D.A.;			
RT	"Complete genome sequence of the model actinomycete Streptomyces			
RT	coelicolor A3(2)."			
RL	Nature 417:141-147(2002).			
DR	Nature: AL939115; CAC09556.1.;			
DR	InterPro: IPR000087; Collagen.			
DR	InterPro: IPR002965; P-rich extensn.			
DR	InterPro: IPR003979; Tropoelastin.			
DR	PRINTS: PR01217; PRODELASTIN.			
DR	PRINTS: PR01500; TROPDELASTIN.			
KW	Hypothetical protein; Complete proteome.			
SEQUENCE	889 AA; 88030 MW; 3AB5F08B1E3E4EF2 CRC64;			
Alignment Scores:				
pred. No.: .	6.08e-06	Length:	889	
Score:	189.50	Matches:	88	
Percent Similarity:	34.13%	Conservative:	12	
Best Local Similarity:	30.03%	Mismatches:	107	
Query Match:	13.56%	Indels:	86	
DB:	16	Gaps:	12	

```

QY 249 CATGCGAGCCCGCGGCGGCTCTCTCCGCGCCGAGGCGCTTGTGCTCCCTGCGCC 190
DB 322 roProGlyPro-----ProProProGlyProProPro---ProGlyProAlaProP 338
QY 189 GGGAGCCTCTGCGCGCGCGCGGCGACCGGCTCTCTGAGGCGCGCAGCATTTGCCCCCTGAGC 130
DB 338 roGlyAlaArgProProProProGlyProProProProGlyProProProProGlyProAlaP 358
QY 129 CATCAGGAATGCCAGGCGCTCTCTGCGCCATCAGCATGCGCGCTCGAACCCCTGTGCCCC 70
DB 358 roProGlyAlaArgProProProProGlyProProProProGlyProProProProGlyProA 378
QY 69 GGCCTTCGCGCTGCAATGCTCCG-----GAGCCTCTGCGCGCGCTCTCAGAGAGAGATC 16
DB 378 laProProGlyAlaArgProProProProGlyProProProProGlyProProProProGlyP 398
QY 15 AGGCGCCACGAGGA 2
DB 398 roAlaProProGly 402

RESULT 10
Q9BIT7 PRELIMINARY; PRT; 1953 AA.
ID Q9BIT7;
AC Q9BIT7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Major amputate spideroin 2-like protein (Fragment).
OS Nephila madagascariensis
OC Araneomorphae; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Nephila;
OC Araneomorphae; Enteretegynae; Araneidae; Tetragnathidae; Nephila.
OX NCBI_Taxid=115969;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21179804; Pubmed=1183372;
RA Gately J., Hayashi C., Motriuk D., Woods J., Lewis R.;
RT "Extreme Diversity, Conservation, and Convergence of Spider Silk
RT Fibrin Sequences."
RL Science 291:2603-2605(2001).
DR EMBL; AF350276; AAK30605.1; -.
FT NON TER 1
FT NON TER 1
FT NON TER 1
SQ SEQUENCE 1953 AA; 159383 MW; C82B4DACD043BD CRC64;

Alignment Scores:
Pred. No.: 1.11e-06 Length: 1953
Score: 199.50 Matches: 66
Percent Similarity: 41.95% Conservative: 7
Best Local Similarity: 37.93% Mismatches: 62
Query Match: 14.22% Indels: 39
DB: 5 Gaps: 7

US-10-023-182-1 (1-752) x Q9BIT7 (1-1953)
QY 3 CCTGTGGGCGCTGACCTTCTCTCTGAGAGCCGCGGAGAGGCTCCGAGCCATGAGGCC 62
DB 1476 ProGlyGlyTyrglyProGlyGlnGlnGlyProGlyGlySerGlyAlaAlaAlaAla 1495
QY 63 GAGAGCGCGGCGACAGGCGGCTGAGCGGCGATGCTGATGCGGCGAGAGGCGCTGAGCAT 122
DB 1496 AlGlyArgGlyProGlyGlyTyrglyProGlyGlnGlnGlyProGlyGlyProGlyAla 1515
QY 123 CCT-----GATGCCACAGGCGGC-----AATGCT 146
DB 1516 AlaAlaAlaAlaAlaAlaGlyThrglyProGlyGlyTyrglyProGlyGlnGlnGlyPro 1535
QY 147 GGGGCGCCAGAGAGAGGCGGCTGCGACGCGCGGAGAGGTCCTCCGCGGCGAGAGG----- 200
DB 1536 GlyGlySerGlyAlaAlaAlaAlaAlaAlaGlyArgGlyProGlyGlyTyrglyProGly 1555
QY 201 -----GCAGCAAGGCGCTCGGCGCGCGAGAGGAGGCGCC 233

```

```

DB 1556 GlnGlnGlyProGlyGlyProGlyAlaAlaAlaAlaAlaGlyProGlyGlyTyrgly 1575
QY 234 CGG-----CGGGGTCCGATGCGCGCGGCTTCA-----GAGCTG 269
DB 1576 ProGlyGlnGlnGlyProGlyAlaAlaAlaAlaAlaAlaGlySerGlyProGlyGly 1595
QY 270 AATGATGCTGAGATGCGCGCGCGAGGCGCGAGAGCGGCTGCTGATTCATCC 329
DB 1596 Tyrgly-ProGlyGlnGlnGlyProGlyGlyProGlyAlaAlaAlaAlaAla----- 1613
QY 330 GCCATGCTTGGCGGACCGCATGAGAGAGGCTGCGCGGAGAGGCTGCGCCAGAT 389
DB 1614 -----GlyArgGlyPro-GlyGly---TyrglyProGlyGlnGlnGlyProGlyG 1629
QY 390 GCGCCACCGCTTCCGTCGCGAGGAGTCTTGAGAGA 427
DB 1629 lyProGlyAlaAlaAlaAlaAlaAlaAlaAlaGlyArgGly 1641

RESULT 11
Q9BX86 PRELIMINARY; PRT; 551 AA.
ID Q9BX86;
AC Q9BX86;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE HPRIT-4 mRNA homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC SPMRAIN-C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; Pubmed=12466851;
RA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team.
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK048615; BAC33392.1; -.
SQ SEQUENCE 551 AA; 59138 MW; A6E14E27DE758CF2 CRC64;

Alignment Scores:
Pred. No.: 3.41e-06 Length: 551
Score: 193.00 Matches: 60
Percent Similarity: 39.87% Conservative: 3
Best Local Similarity: 37.97% Mismatches: 46
Query Match: 11.81% Indels: 49
DB: 11 Gaps: 7

US-10-023-182-1 (1-752) x Q9BX86 (1-551)
QY 412 CCTGCAAGGAGAGCGGTGGGATCCTCGGCGAGGCTCTCGGCGGAGCTGTCTCC 353
DB 208 ProGlyAlaValProGlyGlyAspArgPhe-ProGlyProAlaGlyPro----- 223
QY 352 ATGGGTGTGGAGAGGATGCGAGGTAGAACTCAAGCAGGCGCTCTCCGCGCCCTG 293
DB 224 -----GlyGlyProPro-ProProP 230
QY 292 GCGCGCATCTGACATTCATTCAGGCTGAGAGCCGCGCG-----CCATGCGAGACC 239
DB 230 heProAlaGly-GlnThrProProArgProProLeuGlyProProGlyProProGlyPro 249
QY 238 CGCGGCGCGCTCTCTCCGCGC-----CCGAGGCGCTTGTGCTGCGCGCGCCCGG 188
DB 250 ProGlyProProProProGlyGlnValLeuProProProLeuAlaGlyProProAsnArg 269
QY 187 GGA-----CCTCTG-----CGCCCGGTG 170
DB 270 GlyAspArgProProProProValLeuPheProGlyGlnProPheGlyGlnProProVal 289
QY 169 GCACCGCGCTCTCTGCGCGCGCGAGCATGCGCCCTGCGCGCATCAGGAATGCCAGGCGCT 110

```

```

QY 370 CGGCGCAGCTCTGCTTCATGAGTGTGCGAAGGATGCGAGGTAGAACTCAGCAGG 311
DB 455 oglyProAla-----Pro-----ProGlyAlaArgProPro 466
QY 310 CGGCTCTCCGGCCCCCTTGGCCCGCATCTGCAGCATCATTCACCCCTGAAAGCGCGCCG 251
DB 466 oglyProPro-ProProGlyProProProGly-ProAlaProProGlyAlaArgPro 485
QY 250 CCATCGGAGACCCCGCGGGGCGGCTCTCTCCGAGCCCGGAGCCCTGTGCTGCGCCG 191
DB 466 ProProGlyPro-----ProProProGlyProProPro-----ProGlyProAlaPro 501
QY 190 CGGGGACCTTGGCCCGCGGTGAGCACCCGCTCTCTGGGCGCGCAGCATTTGCCCTGGG 131
DB 502 ProGlyAlaArgProProProGlyProProProProProProProGlyProAla 521
QY 130 CCATGAGGATGCGAGGCGCTCTCTGAGCATGAGCATGCGCCGTCGAACCCCTGTGCGCC 71
DB 522 ProProGlyAlaArgProProProGlyProProProProProProProProGlyPro 541
QY 70 CGGCTTGGCGCTGATGAGCTCCG-----GAGCTCTGCGCGGCTCTCAGAGAGAGGT 17
DB 542 AlaProProGlyAlaArgProProProGlyProProProProGlyProProProGly 561
QY 16 CAGGCGCCAGAGGA 2
DB 562 ProAlaProProGly 566

```

RESULT 8

```

Q95JDO PRELIMINARY; PRT: 511 AA.
ID Q95JDO:
AC Q95JDO: 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Basic proline-rich protein.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Parotid gland;
RA Zhang Q., Szalay A.A., Kyeyune-Nyombi E., Sands J.F., Oberg K.C.,
RA Tieche J.-M., Leonora J.;
RT "Cloning and expression of a novel proline-rich protein from porcine
RT parotid glands.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY035848; AAK61382.1;
DR InterPro; IPR003882; Pstfll_extensin.
DR InterPro; IPR002965; P_rich_extensn.
DR PRINTS; PRO1217; PRICHEXTENSIN.
DR PRINTS; PRO1218; PSTEXTENSIN.
SQ SEQUENCE 511 AA; 48483 MW; AB04597964C448D7 CRC64;

```

Alignment Scores:

```

Pred. No.: 1,94e-07 Length: 511
Score: 210.00 Matches: 59
Percent Similarity: 40.41% Conservative: 0
Best Local Similarity: 40.41% Mismatches: 71
Query Match: 15.02% Indels: 16
DB: 6 Gaps: 5

```

US-10-023-182-1 (1-752) x Q95JDO (1-511)

```

QY 426 CCTTGAAGACACCCCTG---GCACGGAGAGCGGTGGGCGATCTGGCGCAGGCTCTGCG 370
DB 270 ProGlyProAlaProProGlyAlaArgProProProGlyProProProProProPro 289
QY 369 GGGCGAGCTCTGCTTCATGAGGTGTCGAAAGGATGCGAGGTAGAACTCAGCAGGC 310
DB 290 ProProGlyProAlaPro-----HisGlyAlaArgProProPro 302

```

```

QY 309 GGGCTCGGCCCCCTTGGCCCGCATCTGCAGCATTCAGCCCTGAAAGCGCGCCG 250
DB 303 GlyProPro-ProProGlyProProProProGly-ProAlaProProGlyAlaArgPro 322
QY 249 CATTGGAGCCCGCGGGGCGGCTCTCTCCGAGCCCGGAGCCCTTGTCTGCGCCG 190
DB 322 roProGlyPro-----ProProProGlyProProPro-----ProGlyProAlaPro 338
QY 189 GGGAGCTTGGCCCGCGGTGAGCACCCGCTCTCTGGGCGCGCAGCATTTGCCCTGGG 130
DB 338 roGlyAlaArgProProProGlyProProProProProProProProGlyProAla 358
QY 129 CATCAGGATGCGAGGCGCTCTGAGGCGCATGAGCATGCGCCGTCGAACCCCTGTGCGCC 70
DB 358 roProGlyAlaArgProProProGlyProProProProProProProProGlyPro 378
QY 69 GGGCTTGGCGCTGATGAGCTCCG-----GAGCTCTGCGCGGCTCTCAGAGAGAGGT 16
DB 378 laProProGlyAlaArgProProProGlyProProProProGlyProProProGly 398
QY 15 AGGCGCCAGAGGA 2
DB 398 roAlaProProGly 402

```

RESULT 9

```

Q95JDI PRELIMINARY; PRT: 566 AA.
ID Q95JDI:
AC Q95JDI: 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Basic proline-rich protein.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Parotid gland;
RA Zhang Q., Szalay A.A., Kyeyune-Nyombi E., Sands J.F., Oberg K.C.,
RA Tieche J.-M., Leonora J.;
RT "Cloning and expression of a novel proline-rich protein from porcine
RT parotid glands.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY035847; AAK61381.1;
DR InterPro; IPR003882; Pstfll_extensin.
DR InterPro; IPR002965; P_rich_extensn.
DR PRINTS; PRO1217; PRICHEXTENSIN.
DR PRINTS; PRO1218; PSTEXTENSIN.
SQ SEQUENCE 566 AA; 53213 MW; E33B3B5E1BDEB81A CRC64;

```

Alignment Scores:

```

Pred. No.: 1,94e-07 Length: 566
Score: 210.00 Matches: 59
Percent Similarity: 40.41% Conservative: 0
Best Local Similarity: 40.41% Mismatches: 71
Query Match: 15.02% Indels: 16
DB: 6 Gaps: 5

```

US-10-023-182-1 (1-752) x Q95JDI (1-566)

```

QY 426 CCTTGAAGACACCCCTG---GCACGGAGAGCGGTGGGCGATCTGGCGCAGGCTCTGCG 370
DB 270 ProGlyProAlaProProGlyAlaArgProProProGlyProProProProProPro 289
QY 369 GGGCGAGCTCTGCTTCATGAGGTGTCGAAAGGATGCGAGGTAGAACTCAGCAGGC 310
DB 290 ProProGlyProAlaPro-----HisGlyAlaArgProProPro 302
QY 309 GGGCTCGGCCCCCTTGGCCCGCATCTGCAGCATTCAGCCCTGAAAGCGCGCCG 250
DB 303 GlyProPro-ProProGlyProProProProGly-ProAlaProProGlyAlaArgPro 322

```

Query Match: 30.86% Indels: 0
DB: 4 Gaps: 0

US-10-023-182-1 (1-752) x 095987 (1-109)

QY 94 ATGCTGATGAGCCAGAGAGCCCTGGCATTCCTGATGAGCCAGAGGAGCAATGCTGCGGCC 153
| | | | |
DB 1 MetleuMetAlaGlnGlnAlaLeuAlaPheLeuMetAlaGlnGlnAlaMetleuAlaAla 20
QY 154 CAGAGAGAGCGGGTCCAGAGGCGGAGAGGTCCCGGGGCGCAGAGGCGAGCAAGGCGCT 213
| | | | |
DB 21 GlnGlnArgArgValProArgAlaAlaGlnValProGlnAlaGlnGlnGlnGlnPro 40
QY 214 CGGGGCGGGGAGAGAGCGCCCGCGGGGTCCGCATGAGCGGCGGCGCTTCAGGGCATG 273
| | | | |
DB 41 ArgGlnArgGlnGlnAlaProArgGlnValAlaPheMetAlaValProLeuLeuArgArgMet 60
QY 274 GATGCTGAGATGCGGGGCGCAGAGGCGCGAGAGCGCGCTTGAATTCCTGCGCA 333
| | | | |
DB 61 GlnGlnAlaProAlaGlnProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 80
QY 334 TGGCTTTGGCGACACCGCATGAGAGAGAGTGGCGCGGAGAGAGCGCTGCGCGAGATGCC 393
| | | | |
DB 81 CysLeuSerArgArgProTrpLysArgSerTrpSerAlaGlnSerCysProGlnMetPro 100
QY 394 CACCGCTCCCGTCCGAGGCGGTGCTTC 420
| | | | |
DB 101 HisLeuSerProAlaSerProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 109

RESULT 6

Q95146 PRELIMINARY; PRT; 58 AA.

AC 095146; (1-752) (1-109)

DT 01-MAY-1999 (TRENBLREL. 10, Created)

DT 01-MAY-1999 (TRENBLREL. 10, Last sequence update)

DE 01-MAR-2001 (TRENBLREL. 16, Last annotation update)

DE LAGE-2ALT protein isoform.

GN LAGE-2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Wang R.F., Johnston S.L., Zeng G., Topalian S.L.,

RA Schwartzentruber D.J., Rosenberg S.A.,

RT "A Breast and Melanoma-Shared Tumor Antigen: T Cell Responses to

RT Antigenic Peptides Translated from Different Open Reading Frames."

RT J. Immunol. 161:3596-3606(1998).

RN [2]

RP SEQUENCE FROM N.A.

RA Lethe B.G.,

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA MEDLINE-99454989; PubMed-10523621;

RA De Smet C., Lutzquin C., Lethe B., Martelange V., Boon T.,

RT "DNA methylation is the primary silencing mechanism for a set of germ

RT line- and tumor-specific genes with a CpG-rich promoter."

RT Mol. Cell. Biol. 19:7327-7335(1999).

RL EMBL; AJ275977; CAB76944.1; -

DR EMBL; AJ275977; CAB76944.1; -

SQ SEQUENCE 58 AA; 6188 MW; ED12057564BC7EF2 CRC64;

Alignment Scores:

Pred. No.: 6.57e-13 Length: 58
Score: 285.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.31% Indels: 0
DB: 4 Gaps: 0

US-10-023-182-1 (1-752) x 095146 (1-58)

QY 94 ATGCTGATGAGCCAGAGAGCCCTGGCATTCCTGATGAGCCAGAGGAGCAATGCTGCGGCC 153
| | | | |
DB 1 MetleuMetAlaGlnGlnAlaLeuAlaPheLeuMetAlaGlnGlnAlaMetleuAlaAla 20

QY 154 CAGAGAGAGCGGGTCCAGAGGCGGAGAGGTCCCGGGGCGCAGAGGCGAGCAAGGCGCT 213
| | | | |
DB 21 GlnGlnArgArgValProArgAlaAlaGlnValProGlnAlaGlnGlnGlnGlnPro 40

QY 214 CGGGGCGGGGAGAGAGCGCCCGCGGGGTCCGCATGAGCGGCGGCGCTTCAGGGC 267
| | | | |
DB 41 ArgGlnArgGlnGlnAlaProArgGlnValAlaPheMetAlaValProLeuLeuArgArgMet 58

RESULT 7

Q951C9 PRELIMINARY; PRT; 676 AA.

AC 0951C9; (1-752) (1-109)

DT 01-DEC-2001 (TRENBLREL. 19, Created)

DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)

DE 01-MAR-2003 (TRENBLREL. 23, Last annotation update)

DE Basic proline-rich protein.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

OX NCBI_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RA Zhang Q., Szalay A.A., Kyeunye-Nyombi E., Sands J.F., Oberg K.C.,

RA "Cloning and expression of a novel proline-rich protein from porcine

RT parotid glands."

RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY035849; AAK61383.1; -

DR InterPro; IPR005404; KV3.3 channel.

DR InterPro; IPR002965; P rich extensin.

DR PRINTS; PR01582; KV33CHANNEL.

DR PRINTS; PR01217; PRICEXTENSIN.

SQ SEQUENCE 676 AA; 62297 MW; 3008BC41EFD81FC9 CRC64;

Alignment Scores:

Pred. No.: 1.07e-07 Length: 676
Score: 213.50 Matches: 82
Percent Similarity: 34.82% Conservative: 4
Best Local Similarity: 33.20% Mismatches: 101
Query Match: 15.27% Indels: 60
DB: 6 Gaps: 11

US-10-023-182-1 (1-752) x 0951C9 (1-676)

QY 709 AGCGTCTGCTGCTGAGAGCAAAACATGAGCCGCCCAATGAGTGGCACTGCTGCTG 650
| | | | |
DB 371 AlaArgProProProProProProProProProProProProProProProProProPro 385

QY 649 GAACCATTCCTTGAAGGAGAGA-----GGCATACCTTAGAGAGG-----GCG 608
| | | | |
DB 386 GlnProAlaProProProGlnAlaArgProProProProProProProProProProProPro 405

QY 607 CGAGGCTGGCTTGAAGGCGCTGCTGAGAGGAGGAGGAGCAAAACAGCGGAGAGAG 548
| | | | |
DB 406 ProGlnProAla-----ProProGlnAlaArgProProProProProProProProProPro 421

QY 547 CACTGCGTATTCACATCAACAGGAAAGCT---GCTGAGACAGAGAGCTGATGAGAGC 491
| | | | |
DB 421 oProAlaAspGlnProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 436

QY 490 TGGAGTTGGGGGAGGCTGTGAGACAGTCAGTCGATGATGATGATGATGATGATGATGATG 431
| | | | |
DB 436 ----- 436

QY 430 AACTCTTCAGAGAACCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 371
| | | | |
DB 437 -----ProLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 455

RA Nelson D.L.;
 RT "Multiple pathogenic and benign genomic rearrangements occur at a 35-
 kb duplication involving the NEMO and the IKG2 genes."
 RL Hum. Mol. Genet. 0:0-0(2001).
 DR EMBL; AJ012834; CAA10194.1; -
 DR EMBL; AF277315; AAL27015.1; -
 SQ SEQUENCE 180 AA; 18236 MW; 9077PAF953543A25 CRC64;

Alignment Scores:

Pred. No.:	1,166-51	Length:	180
Score:	814.00	Matches:	152
Percent Similarity:	91.67%	Conservative:	13
Best Local Similarity:	84.44%	Mismatches:	15
Query Match:	58.02%	Indels:	0
DB:	4	Gaps:	0

US-10-023-182-1 (1-752) x Q9Y479 (1-180)

```

QY 54 ATGACAGCCGAAAGCCGCGGACAGAGGCGTTCGACGGCGATGCTGATGCCAGAGGC 113
DB 1 MetGlnAlaGlnGlyGlnGlyThrGlySerThrGlyAspAlaAspGlyProGlyGly 20
QY 114 CCGGCAATTCCTGATGCGCCAGAGGCGCAATGCTGCGGCGCCAGAGAGCGGCGCCAG 173
DB 21 ProGlyIleProAspGlyProGlyGlnAlaGlyGlyProGlyGlnAlaGlyAlaThr 40
QY 174 GCGCGCAGAGTCCCGCGGCGCAGAGGCGCAGAGGCGCTCGGCGCCGCGAGAGCGCC 233
DB 41 GlyIYArgGlyProArgGlyYalaglyAlaAlaArgAlaSerGlyProArgGlyYalYala 60
QY 234 CCGCGGCGGCGCCAGAGGCGCGCGCGCGCTTCAGAGGCTGAATGATGTCAGATGCGGCGCC 293
DB 61 ProArgGlyProIleGlyYalAlaAlaSerAlaGlnAlaSerGlyArgGlyProCysGlyYala 80
QY 294 AGGCGGCGCGAGAGCGCGCTGCTGATGCTACCTCGCCATGCTCTTGCGCAGACCCATG 353
DB 81 ArgArgProAspSerArgIleuGlnIleuHisIleThrMetProPheSerSerProMet 100
QY 354 GAGCGAGAGTCCCGCGCAGAGGCGCTGCGCCAGATGCGCCAGCGCTTCCGTCGAGG 413
DB 101 GlnAlaGlnIleuValArgArgIleuSerArgAspAlaAlaProIleuProArgProGly 120
QY 414 GTGCTTCGAGAGAGTTCAGTGTCTCGGCAACATACATGACTATCCGACTGCTGCA 473
DB 121 AlaValIleuLysAspPheThrValSerGlyAsnIleuPheIleArgIleuThrAlaAla 140
QY 474 GACCAACCGCACTGCAAGCTCTTCATCAGCTCTGCTTCAGAGCTTCCCTGTTGATG 533
DB 141 AspHisArgGlnIleuGlnIleuSerIleSerCysIleuGlnIleuSerIleuMet 160
QY 534 TGGATCAGCAGAGTCTTCTGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 593
DB 161 TrpIleThrGlnCysPheIleuProValPheIleuAlaGlnAlaProSerGlyGlnArgArg 180

```

RESULT 2

Q9BU80 PRELIMINARY; PRT; 210 AA.

ID Q9BU80
 AC Q9BU80;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Cancer/Leucis antigen 2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA Strausberg R.;
 RL EMBL; BC02833; AAH02833.1; -
 DR EMBL; BC02833; AAH02833.1; -
 SQ SEQUENCE 210 AA; 21089 MW; 8F55BF04FB04E8BE CRC64;

Alignment Scores:

Pred. No.:	7,636-36	Length:	210
Score:	598.00	Matches:	121
Percent Similarity:	68.12%	Conservative:	20
Best Local Similarity:	58.45%	Mismatches:	38
Query Match:	42.62%	Indels:	28
DB:	4	Gaps:	4

US-10-023-182-1 (1-752) x Q9BU80 (1-210)

```

QY 54 ATGACAGCCGAAAGCCGCGGACAGAGGCGTTCGACGGCGATGCTGATGCCAGAGGC 113
DB 1 MetGlnAlaGlnGlyArgGlyThrGlySerThrGlyAspAlaAspGlyProGlyGly 20
QY 114 CCGGCAATTCCTGATGCGCCAGAGGCGCAATGCTGCGGCGCCAGAGAGCGGCGCCAG 173
DB 21 ProGlyIleProAspGlyProGlyGlnAlaGlyGlyProGlyGlnAlaGlyAlaThr 40
QY 174 GCGCGCAGAGTCCCGCGGCGCAGAGGCGCAGAGGCGCTCGGCGCCGCGAGAGCGCC 233
DB 41 GlyIYArgGlyProArgGlyYalaglyAlaAlaArgAlaSerGlyProArgGlyYalYala 60
QY 234 CCGCGGCGTCCGATGCGCGCGCGCGCTTCAGAGGCTGAATGATGCTGCAATGCGGCGCC 293
DB 61 ProArgGlyProIleGlyYalAlaAlaSerAlaGlnAlaSerGlyArgCysProCysGlyYala 80
QY 294 AGGCGGCGCGAGAGCGCGCTGCTGATGCTACCTCGCCATGCTCTTGCGCAGACCCATG 353
DB 81 ArgArgProAspSerArgIleuGlnIleuHisIleThrMetProPheSerSerProMet 100
QY 354 GAGCGAGAGTCCCGCGCAGAGGCGCTGCGCCAGATGCGCCAGCGCTTCCGTCGAGG 413
DB 101 GlnAlaGlnIleuValArgArgIleuSerArgAspAlaAlaProIleuProArgProGly 120
QY 414 GTGCTTCGAGAGAGTTCAGTGTCTCGGCAACATACATGACTATCCGACTGCTGCA 473
DB 121 AlaValIleuLysAspPheThrValSerGlyAsnIleuPheIleArgIleuThrAlaAla 140
QY 474 GACCAACCGCACTGCAAGCTCTTCATCAGCTCTGCTTCAGAGCTTCCCTGTTGATG 533
DB 141 AspArgGlnGly-----AlaGlyArgMetArgYalYalGly 152
QY 534 TGGATCAGCAGAGTCTTCTGCGCGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 590
DB 153 Trp-----GlyIleuGlySerAlaSerProGlnGlyIleuYala 165
QY 591 -----CGCTAAGCCAGCGCTGCGCGCTTCTGCTG 620
DB 166 ArgAspIleuArgThrProLysIleuValSerGlnIleuArgProGlyThrProGly--- 184
QY 621 GTGATGCTTCCTGCGCGCTTCTGCGCGCTTCTGCGCGCTTCTGCGCGCTTCTGCG 641
DB 185 ---ProProProProGlnGly 190

```

RESULT 3

Q9U89 PRELIMINARY; PRT; 210 AA.

ID Q9U89
 AC Q9U89;
 DT 01-MAY-2000 (T-EMBlrel. 13, Created)
 DT 01-MAY-2000 (T-EMBlrel. 13, Last sequence update)
 DT 01-OCT-2002 (T-EMBlrel. 22, Last annotation update)
 DE LAGE-1L protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Melanoma;
 RX MEDLINE=99325550; PubMed=10399963;
 RA Aarnoudse C.A.; Van den Doel P.B.; Heemskerk B.; Schrier P.I.;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 10, 2004, 17:29:51; Search time 42.5 Seconds
(without alignments)
9132.026 Million cell updates/sec

Title: US-10-023-182-1
Perfect score: 1403
Sequence: 1 ATCTCTGTGGCCCTGACCT.....TAAACTAGCTACGAAAA 752

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 1661050

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ n2p.model -DEV=xip
-Q/cgcn2.1/USFTO.spool.p/US10023182/runat_10022004.171429.3877/app_query.fasta_1.903
-DB=SPTRMBL_23 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bites -START=1 -END=1 -MATRIX=bl6sum62 -TRANS=human40.cdt -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRF=US10023182 @CGN 1.1.86 @runat_10022004.171429.3877 -NGPU=6 -ICPU=3
-NO MMAP -LARGEJOLIB -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -TREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: SPTRMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp Unclassified:*
15: sp Yivirus:*
16: sp Bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	814	58.0	180	4	Q9Y479 homo sapien

2	598	42.6	210	4	Q9BU80	Q9BU80 homo sapien
3	591	42.1	210	4	Q9YU89	Q9YU89 homo sapien
4	581	41.4	142	4	Q9NY13	Q9NY13 homo sapien
5	433	30.9	109	4	Q9Y987	Q9Y987 homo sapien
6	285	20.3	58	4	Q9Y146	Q9Y146 homo sapien
7	213.5	15.3	676	6	Q9YJC9	Q9YJC9 sus scrofa
8	210	15.0	511	6	Q9YJDO	Q9YJDO sus scrofa
9	210	15.0	566	6	Q9YJDI	Q9YJDI sus scrofa
10	199.5	14.2	1953	5	Q9YBIT	Q9YBIT neophila mad
11	193	13.8	551	11	Q9BX86	Q9BX86 mus musculus
12	190	13.6	889	16	Q16630	Q16630 homo sapien
13	189.5	13.6	889	16	Q9FZNS	Q9FZNS streptomyce
14	188.5	13.5	922	5	Q8MW53	Q8MW53 mytilus gal
15	188.5	13.4	853	2	Q8VQZ0	Q8VQZ0 myxococcus
16	188.5	13.4	857	2	Q8Y783	Q8Y783 myxococcus
17	188.5	13.4	905	5	Q8MW55	Q8MW55 mytilus gal
18	186.5	13.3	1215	5	Q77202	Q77202 acanthamoeb
19	186	13.3	304	5	Q20739	Q20739 caenorhabdi
20	185.5	13.3	520	11	Q61078	Q61078 mus musculus
21	185.5	13.2	604	16	Q9L252	Q9L252 streptomyce
22	185.5	13.3	1461	4	Q76045	Q76045 homo sapien
23	185.5	13.3	1464	4	Q8N473	Q8N473 homo sapien
24	185	13.2	682	5	Q9VRM2	Q9VRM2 drosophila
25	184.5	13.2	525	10	Q8W158	Q8W158 brassica ol
26	184.5	13.2	922	5	Q4367	Q4367 mytilus edu
27	184.5	13.2	1002	5	Q9BIU8	Q9BIU8 argiope tri
28	183.5	13.1	302	5	Q19079	Q19079 caenorhabdi
29	183	13.1	783	16	Q9XAI1	Q9XAI1 streptomyce
30	183	13.1	1453	11	Q63079	Q63079 rattus norv
31	182	13.0	410	5	Q16988	Q16988 araneus dia
32	181.5	13.0	268	5	Q8T018	Q8T018 drosophila
33	181	12.9	1431	11	Q9JMH4	Q9JMH4 mesocricetu
34	181	12.9	3247	12	Q65553	Q65553 bovine herp
35	180.5	12.9	2249	5	Q9NMH4	Q9NMH4 neophila cla
36	180	12.9	551	11	Q8BX18	Q8BX18 mus musculus
37	180	12.8	1884	5	Q9NMW2	Q9NMW2 neophila mad
38	179.5	12.8	420	5	Q9VZC2	Q9VZC2 drosophila
39	179	12.8	389	11	Q88539	Q88539 mus musculus
40	179	12.8	641	11	Q923D5	Q923D5 mus musculus
41	178.5	12.8	409	10	Q9SBM1	Q9SBM1 volvox cart
42	178.5	12.8	641	11	Q8VD10	Q8VD10 mus musculus
43	178.5	12.8	1497	4	Q9NMK9	Q9NMK9 homo sapien
44	178.5	12.8	1497	4	Q9UMD9	Q9UMD9 homo sapien
45	178.5	12.8	1532	4	Q02802	Q02802 homo sapien

ALIGNMENTS

RESULT 1

ID	Q9Y479	PRELIMINARY	PRT	180 AA.
AC	Q9Y479			
AD	Q9Y479			
DT	01-NOV-1999 (TRENBLrel. 12, Created)			
DT	01-NOV-1999 (TRENBLrel. 12, Last sequence update)			
DT	01-OCT-2002 (TRENBLrel. 22, Last annotation update)			
DE	LAGE-18 protein (Cancer/testis antigen 2).			
GN	LAGE1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Melanoma;			
RX	MEDLINE=9325550; PubMed=1039963; Heemkerk B, Schrier P.I.,			
RA	Aarrouse C.A., Van den Doel P.B., Heemkerk B, Schrier P.I.,			
RT	"Interleukin-2-induced, melanoma-specific T cells recognize CMEL, an			
RI	unexpected translation product of LAGE-1."			
RI	Int. J. Cancer 82:442-448(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Arachya S., Bardeiro T., Galgoczy P., Yamagata T., Esposito T.,			
RA	Patlan H., Ciccodicola A., Kenwick S., Platzer M., D'Urso M.,			

Search completed: February 10, 2004, 17:38:01
Job time : 30 secs

```
QY 452 GACTATCCAGTACTGCTGCTGCAAGACCCGCACTGCTTCATCAGCTCTGTTCT 511
   |||
Db 771 GYPRoProGlyGluArgGlyAlaProGlyProGln-----Gly 783
QY 512 CCAGCAGCTTCCCTGTGATGTGATCAAGCAGTGTCT--GCCGTTGTTTGGC 568
   |||
Db 784 ProProGlyAlaProGlyProLeuGlyIleAlaGlyLeuThrGlyAlaArg-GlyLeuAl 803
QY 569 TCAGCTCCCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 628
   |||
Db 803 aglyProPro-----GlyMetProGlyAlaArgGly-----SerPr 815
QY 629 TCCTCCCTTGGG-----AATGT-----CCAGCAGGAGTGGCCAG 665
   |||
Db 815 OGlyProGlnGlyIleArgGlyGluGlyGlyProGlyProSerGlyGln 832
```

RESULT 15

A24264

proline-rich protein MP2 - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 09-Sep-1987 #sequence_revision 09-Sep-1987 #text_change 03-May-1996

C:Accession: A24264

R:Am, D.K.; Carlson, D.M.

J. Biol. Chem. 260, 15863-15872, 1985

A:Title: The structure and organization of a proline-rich protein gene of a mouse multi-

A:Reference number: A92508; MUID:86059475; PMID:2999141

A:Accession: A24264

A:Molecule type: DNA

A:Residues: 1-240 <ANN>

C:Superfamily: proline-rich protein

Alignment Scores:

Pred. No.:	2.86e-05	Length:	240
Score:	178.50	Matches:	57
Percent Similarity:	41.89%	Conservative:	55
Best Local Similarity:	38.51%	Mismatches:	31
Query Match:	12.77%	Indels:	9
DB:	2	Gaps:	9

US-10-023-182-1 (1-752) x A24264 (1-240)

```
QY 426 CCTTCAGAACACCCCTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 367
   |||
Db 61 ProProGlyGlyProGlnProArgProProGlnGlyProProProProGly-----Gly 78
QY 366 CCAGCTCTGCTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 307
   |||
Db 79 ProGlnProArgPro-----ProGlnGly 86
QY 306 TCTCCGAGCCCTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 247
   |||
Db 87 ProPro-ProProGlyGlyProGln-GlnArgPro-----ProGlnGlyProProPro 104
QY 246 GCGGA-----CCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 205
   |||
Db 104 roGlyGlyProGlnProArgProProGlnGlyProProProProGlnGlyProGlnLeuAl 124
QY 204 CTGCCCCCTGCGCCCGGAGCTCTGCGCCCTGCGCCCTGCGCCCTCTCTGCGCCGCGAG 145
   |||
Db 124 rg-----ProProGlnGlyProProProProAlaGlyPro--GlnProArgProPro 141
QY 144 CATTGCCCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 85
   |||
Db 141 lnglyProProProProAlaGly--ProGlnProArgProProGlnGlyProProThr 160
QY 84 AACCCCTGTGCCCCGAGCTTGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 28
   |||
Db 160 hrglyProGlnProArgProThr-----GlnGlyProProProProGlnGlyProGlnG 178
QY 27 AGAGAGAGGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8
   |||
Db 178 lngArgProProGlnGlyPro 184
```



```

Db      127 ylyseProboValala-ProCygluProthrProProCyslyseProCysProG 147
Qy      237 GCGGGGCGGCT---CCTCCCGGCCCC-----215
Qy      147 lnglyProProglProProglProProglProglProglProglProglProgl 167
Qy      214 -----GAGGCCCTTCTGCTCCCTCCGCCCCCGGAGGACTCTGC 178
Db      167 hrProglYArProglYlThraPaAlaProglYserProglYProArGlYlProProG 187
Qy      177 CGCGCGGTGACCGCCGCTC-----TCTCCGCGCGCGCGACATG- 140
Db      187 lProAlaGlYlAlaGlYAlaProglYProAlaGlYlProAlaGlYlThraProAla 207
Qy      139 -----CCCCCTGGGCGCATCGAAGATGCGAGGCGCTC 109
Db      207 erGlUProleuthrProglYAlaProglYglUProglYAspSerGlYProProGlYProP 227
Qy      108 CTGGGCGCATCGACATCGCCGCGCA-----CCCCGTGCGCCGCGCTCGGCT 58
Db      227 roGlYProProglYAlaProglYAlaProglYAspSerGlYProProGlYProlyse 247
Qy      57 GCATGCTCCGAGCGCTCTCCCGGCTCTCAGAGAGAGTCA--GGCCGACAGCA 2
Db      247 laProGlYProAspGlYProProglYAla-----AspGlYlSerGlYProProGlY 264

```

RESULT 14

```

CSB07S
collagen alpha 1(III) chain - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 04-Dec-1986 #sequence revision 04-Dec-1986 #text change 07-May-1999
C:Accession: A02862; A38001; A38002; A38003; A38004; A38005; S71946
R:Fietzek, P.P.; Allmann, H.; Rauterberg, J.; Henkel, W.; Wächter, E.; Kuehn, K.
Hoppe-Seyler's Z. Physiol. Chem. 360, 809-820, 1979
A>Title: The covalent structure of calf skin type III collagen. I. The amino acid sequen
A:Reference number: A02862; PMID:80026026; PMID:488906
A:Accession: A02862
A:Molecule type: protein
A:Residues: 1-242 <FIE>
R:Dewes, H.; Fietzek, P.P.; Kuehn, K.
Hoppe-Seyler's Z. Physiol. Chem. 360, 821-832, 1979
A>Title: The covalent structure of calf skin type III collagen. II. The amino acid sequen
A:Reference number: A38001; PMID:80026027; PMID:488907
A:Accession: A38001
A:Molecule type: protein
A:Residues: 243-422 <DEW1>
R:Bentz, H.; Fietzek, P.P.; Kuehn, K.
Hoppe-Seyler's Z. Physiol. Chem. 360, 833-840, 1979
A>Title: The covalent structure of calf skin type III collagen. III. The amino acid sequen
A:Reference number: A38002; PMID:80026028; PMID:488908
A:Accession: A38002
A:Molecule type: protein
A:Residues: 423-571 <BEN>
R:Lang, H.; Glanville, R.W.; Fietzek, P.P.; Kuehn, K.
Hoppe-Seyler's Z. Physiol. Chem. 360, 841-850, 1979
A>Title: The covalent structure of calf skin type III collagen. IV. The amino acid sequen
A:Reference number: A38003; PMID:80026029; PMID:488909
A:Accession: A38003
A:Molecule type: protein
A:Residues: 572-808 <LAN>
R:Dewes, H.; Fietzek, P.P.; Kuehn, K.
Hoppe-Seyler's Z. Physiol. Chem. 360, 851-860, 1979
A>Title: The covalent structure of calf skin type III collagen. V. The amino acid sequen
A:Reference number: A38004; PMID:80026030; PMID:488910
A:Accession: A38004
A:Molecule type: protein
A:Residues: 809-947 <DEW2>
R:Allmann, H.; Fietzek, P.P.; Glanville, R.W.; Kuehn, K.
Hoppe-Seyler's Z. Physiol. Chem. 360, 861-868, 1979
A>Title: The covalent structure of calf skin type III collagen. VI. The amino acid sequen
A:Reference number: A38005; PMID:80026031; PMID:488911
A:Accession: A38005
A:Molecule type: protein

```

```

A:Residues: 948-1049 <ALL>
A:Experimental source: skin
R:Henkel, W.
Biochem. J. 318, 497-503, 1996
A>Title: Cross-link analysis of the C-telopeptide domain from type III collagen.
A:Reference number: S71946; PMID:96404897; PMID:8809038
A:Accession: S71946
A:Molecule type: protein
A:Residues: 87-106;1017-1029;1037-1049 <HEN>
C:Comment: Prolines at the third position of the tripeptide repeating unit (G-X-Y) are
C:Comment: The type III collagen molecule is a trimer of identical chains, linked to ea
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology
C:Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxyprol
F:11049/Product: collagen alpha 1(III) chain #status experimental <CAB>
F:1114/Region: amino-terminal nonhelical telopeptide
F:1151040/Region: helical
F:1587589/Region: cell attachment (R-G-D) motif
F:752754/Region: cell attachment (R-G-D) motif
F:875877/Region: cell attachment (R-G-D) motif
F:878880/Region: cell attachment (R-G-D) motif
F:935937/Region: cell attachment (R-G-D) motif
F:10411049/Region: carboxyl-terminal nonhelical telopeptide
F:95197,119,938,950/Modified site: 5-hydroxylysine (lys) #status experimental
F:107,950/Modified site: allysine (lys) #status predicted
F:107/Binding site: carbohydrate (lys) (covalent) #status experimental
F:1040,1041/Disulfide bonds: interchain #status predicted

```

Alignment Scores:

```

Pred. No.: 2,3e-05 Length: 1049
Score: 179.00 Matches: 79
Percent Similarity: 36.828 Conservative: 16
Best Local Similarity: 30.624 Mismatches: 98
Query Match: 12.764 Indels: 66
DB: 1 Gaps: 12
US-10-023-182-1 (1-752) x CSB07S (1-1049)
Qy      3 CCTGTGGGCGGCGTACCTTCTCTGAGAGCGGCGAGAGGCTCCGAGCCATGAGGCC 62
Db      604 ProArGlYProhrGlYProhrGlYProhrGlYProhrGlYProhrGlYProhrGlY 617
Qy      63 GAGGCGGCGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 122
Db      618 -----GlylnProglYAspSerGlYlSerGlYlSerGlYlSerGlYlSerGlYl 634
Qy      123 -----CCTGATGCGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 140
Db      635 AlAGlYProArGlYlYlProglYlUArGlYlUArGlYlUArGlYlUArGlYlUArGlY 654
Qy      141 AATGCTGGGCGGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 197
Db      655 PheProGlYAlaProglYlUArGlYlUArGlYlUArGlYlUArGlYlUArGlYlUArGlY 674
Qy      198 -----GGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 233
Db      675 GlYlUlyseGlYlUlyGlYlUlyGlYlUlyGlYlUlyGlYlUlyGlYlUlyGlYlUly 694
Qy      234 -----CCGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 272
Db      695 AlAGlYProProglYlProglYlUArGlYlUArGlYlUArGlYlUArGlYlUArGlYl 714
Qy      273 GATGTGTGAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 331
Db      715 AlAlAGlYlYlPheProglYlYlPheProglYlYlPheProglYlYlPheProglYlY 730
Qy      332 CATGCTTGTGCGACACCATGAGACAGAGCTGCGCCGAGAGGAGGAGGAGGAGGAGGAGGAG 391
Db      731 AsnGlYAsnProglYlYlProProglYlYlSerSerGlYlAlaProglYlYlAspSer 750
Qy      392 CCCACGCGTCCCGTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 451
Db      751 ProProGlYlSerAsnGlYlAlaProglYlSerProglYlYlSerGlYlProlyseGlYlAspSer 770

```

A/Cross-references: EMBL:X62379; NID:G51552; PIDN:CAA44244.1; PID:G51553

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	1,176-05	1206	46	7	47	30
Percent Similarity:	40.77%					
Best Local Similarity:	35.38%					
Query Match:	13.09%					
DB:	2					4

US-10-023-182-1 (1-752) x S24407 (1-1206)

```

QY 322 AACTGACGAGCGCGCTCTCCGCGCCCTGCGCCCGCATGAGCATTCATTCAGCCCT 263
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 636 SerSerSerGlnGlnIlySerProProAlaProProThrProProLeuProPro 655
QY 262 GAAGCCGCGCGCCCATGCGGACCCCGCGGCGCCTCTCTCCGCGCCCGCATTCGCT 203
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 656 ProLeuIleProPro-----ProProProLeuProProGlyLeuGly 669
QY 202 GCCCTGCGCGCCCGCGGACCTCTGCGCGCCGCGCATCCGCTCTCTGCGCGCCGCA 143
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 670 ProLeuProProAlaProProIleProProValCysProValSerProProProPro 699
QY 142 TTGCCCCCTGCGCGCATGAGATGCCA-----GCGCTCTCTGCGGCA----- 101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 690 ProProProProProProProProValProProSerAspGlyProProProProPro 709
QY 100 -----TCAGCATGCGCGCGCGCAACC 80
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 710 ProProProProLeuProProAlaLeuAlaLeuProProSerGlyGlyProProProPro 729
QY 79 CCTGTGCCCCCGCTTGGCGCTTGCATGCGCTCGGAGCCTTGGCGGC----- 32
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 730 ProProProProProProProProGlyLeuAlaProProProProProGlyLeuSerPheGly 749
QY 31 ---TCTCAGAGAGAAAGTCAGGCGCCACGA 5
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 750 LeuSerSerSerSerSerGlnTyProArg 759

```

RESULT 12

formin - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 22-Jan-1994 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999
 C/Accession: S11515
 R/Moychik, R.P.; Maas, R.L.; Zeller, R.; Vogt, T.F.; Leder, P.
 Nature 346, 850-853, 1990
 A/Title: 'Formins', proteins deduced from the alternative transcripts of the limb deform
 A/Reference number: S11515; MUID:90363291; PMID:2392150
 A/Accession: S11515
 A/Molecule type: mRNA
 A/Residues: 1-1468 <MOY>
 A/Cross-references: EMBL:X53599; NID:G52877; PIDN:CAA37668.1; PID:G52878

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	1,156-05	1468	46	7	47	30
Percent Similarity:	40.77%					
Best Local Similarity:	35.38%					
Query Match:	13.09%					
DB:	2					4

US-10-023-182-1 (1-752) x S11515 (1-1468)

```

QY 322 AACTGACGAGCGCGCTCTCCGCGCCCTGCGCCCGCATGAGCATTCATTCAGCCCT 263
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 862 SerSerSerGlnGlnIlySerProProAlaProProThrProProProProPro 881
QY 262 GAAGCCGCGCGCCCATGCGGACCCCGCGGCGCCTCTCTCCGCGCCCGCATTCGCT 203
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 882 ProLeuIleProPro-----ProProProLeuProProGlyLeuGly 895

```

```

QY 202 GCCCTGCGCGCCCGCGGACCTCTGCGCGCCGCGCATCCGCTCTCTGCGCGCCGCA 143
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 896 ProLeuProProAlaProProIleProProValCysProValSerProProProPro 915
QY 142 TTGCCCCCTGCGCGCATGAGATGCCA-----GCGCTCTCTGCGGCA----- 101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 916 ProProProProProProProProValProProSerAspGlyProProProProPro 935
QY 100 -----TCAGCATGCGCGCGCGCAACC 80
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 936 ProProProProLeuProProAlaLeuAlaLeuProProSerGlyGlyProProProPro 955
QY 79 CCTGTGCCCCCGCTTGGCGCTTGCATGCGCTCGGAGCCTTGGCGGC----- 32
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 956 ProProProProProProProProGlyLeuAlaProProProProProGlyLeuSerPheGly 975
QY 31 ---TCTCAGAGAGAAAGTCAGGCGCCACGA 5
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 976 LeuSerSerSerSerSerGlnTyProArg 985

```

RESULT 13

collagen 1 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 21-Jul-2000
 C/Accession: A31219; T37290; T37291
 R/Kramer, J.M.; Cox, G.N.; Hlrich, D.
 Cell 30, 599-606, 1982
 A/Title: Comparisons of the complete sequences of two collagen genes from Caenorhabditis
 A/Reference number: A30826; MUID:83050944; PMID:7139711
 A/Accession: A31219
 A/Molecule type: DNA
 A/Residues: 1-236 <KRA>
 A/Cross-references: GB:V00147; NID:G6677; PIDN:CAA23463.1; PID:G6678

A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-236 <KR2>
 A/Cross-references: EMBL:J01047; PIDN:AA27988.1
 A/Accession: T37291
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-236 <KR3>
 A/Cross-references: EMBL:V00147; NID:G6677; PIDN:CAA23463.1; PID:G6678
 C/Genetics:
 A/Gene: col-1
 A/Intons: 54/3; 280/2
 C/Superfamily: unassigned collagens

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	1,866-05	296	66	9	62	64
Percent Similarity:	161.00					
Best Local Similarity:	37.31%					
Query Match:	32.84%					
DB:	2					11

US-10-023-182-1 (1-752) x A31219 (1-296)

```

QY 426 CCTTGAAGAGACCCCTGCGACGAGCGGTGGGCGATCCTGGG----- 382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 70 ProProArgSerArgThrThrArgGlnAlaTyGlyGlyProGlyValAlaSerProAlaPro 89
QY 381 -----CCAGCTCTCTGCGGCGAGCTCTGCTTCATCG 349
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 90 AsnLeuGlnCysGlnGlyCysCysLeuProGlyProProGlyProAlaAlaGlyAlaProGly 109
QY 348 GTGTCCGAAAGGACGAGCGGAGGTAGACTCAAGAGGCGGCTCCG---GCCCGCTGG 292
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 110 LysProGlyLysPro-----GlyArgProGlyAlaProGlyThrProGlyThr-ProGly 127
QY 291 CCGCGATCTGACGATTCATTCAGCCCTGAAGCCGCGCGCGCATGCGGA-----CCCC 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

DB: 2 Gaps: 9

US-10-023-182-1 (1-752) x T15936 (1-302)

```

QY 447 TGTTCGCCGACAGTGAACCTCTTACAGAACACCCCTGGACCGGAAAGCGGTGGGCGAT 388
Db 105 CysGcSThrCysGlnValGlyProProGlyProProGlyProProGlyAArgAspGlyArg 124
QY 387 CCTGGG-----CCAGGCTCCTGGGCGCA-----GCTCTGCTT 355
Db 125 ProGlyAlaProGlyArgProGlyAsnProGlyArgAspGlyArgAlaLeu 144
QY 354 CCATGGTGTGCGAAAGGACATGGGAGTAGAACTCAAGACGAGCGGCTCCGCGCCCG 295
Db 145 Pro-----GlyProProProLeuPro 151
QY 294 TGGCCCCGACATTCAGCATTCATTCAGCCCTGAAGCGCGCCGACATGGCGACCCCG 235
Db 152 ProCysGlnCysPro--ProGlyProProGlyProAlaGlyProProGlyProLeu 171
QY 234 GGGGCGCT----- 227
Db 171 LysProGlyProGlnGlyAspAlaGlyThrSerGlyGlnAspGlyValProGlyLeu 191
QY 226 -----CTTCCGCGCCCGGAC-----GCCCTTGCTGCCCCCTGCGCCCGGAGAC 184
Db 191 roGlyProProGlyProSerGlyProGlnGlyAlaProGlyValProGlyGlnLysGlyP 211
QY 183 CT-----CTGCGCGCGGTGGACACCGGCTCTGCTGGCGCGCGAG 145
Db 211 roThGlyGlnProGlyLysValIleAsnGlyAlaProProGlyProProGlyProProG 231
QY 144 CATTCGCCCTGGGACATCAAGATGCAAGGCGCTCTGCGGACATCAAGATCGCCGCTG 85
Db 231 Ly--ProProGlyProGlnGlyProProGlyProProGlyLysAspGlyGlnProGlyL 250
QY 84 AA-----GCCCTGTGCGCCGCGCTTGGCGCTGATGGCTCGGAGGCTCTGCGG 34
Db 250 yAlaGlyProProGlyLeuProGlyAspProGlyGlnLysGlySerAspGlyLeuProG 270
QY 33 GCTCTCAGAGAGAGTCAAGGCCCAAGAGA 2
Db 270 Ly--ProHisGlyGlyThrGlyProArgGly 279

```

RESULT 10

T35389 Probable serine-threonine protein kinase - Streptomyces coelicolor

C/Species: Streptomyces coelicolor

C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 21-Jan-2000

C/Accession: T35389

R/Murphy L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, June 1999

A/Reference number: Z21576

A/Accession: T35389

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-783 <MUR>

A/Cross-references: EMBL:AL079348; PDB:CB45488.1; GSPDB:GN00070; SCQDB:SC66T3.32C

A/Experimental source: strain A3(2)

A/Genetic: SCQDB:SC66T3.32C

A/Suprafamily: unassigned collagens

Alignment Scores:

```

Pred. No.: 1,22e-05 Length: 783
Score: 183.00 Matches: 85
Percent Similarity: 33.57% Conservative: 10
Best Local Similarity: 30.04% Mismatches: 95
Query Match: 13.04% Indels: 93
DB: 2 Gaps: 12

```

US-10-023-182-1 (1-752) x T35389 (1-783)

```

QY 36 GGCAGAGGCTCCGAGACATGACGCGGAGCGCGGACAGGAGGCTTCAGCGGCGAT 95
Db 373 GlyGlyGlyArgGlyGlyValGlyProGlyGlyAlaGlyProGlyGlyValGlyProGly 392
QY 96 GCTGATGCGCCAGAGGCGCTTGGCATTTCTGATGGCCAGAGGGGCAATGCTGGCGCCA 155
Db 393 GlyValGlyProGlyGlyValGly--ProGlyGlyValGlySerGlyGlyValGlyPro 411
QY 156 GGAAGAGCGGCT-----GCCACGCGCGGACAGAGTCC 188
Db 412 GlyGlyAlaGlyProGlyGlyValGlyProGlyGlyAlaGlySerGlyGlyValGlyPro 431
QY 189 CGGGGCGAGGCGGACGACAGGCGCTCGGCGCGGAGAGAGGCGCGCGGCT---CCG 245
Db 432 GlyGlyAlaGlySerGly--GlyValGlyProGlyGlyAlaGlySerGlyGlyValGly 450
QY 246 CATGGCGGCGGCGCTTCAAGGCTGAATGATGCTGACATGCGGCGGCGGCGGCGGAG 305
Db 451 ProGlyGlyAlaGlySerGlySerGlySerAlaGly--ProAspGlyAlaAspProGlyGly 468
QY 306 AGCGGCTGCTTGAATTTACTGCGCATGCTTTCGCGACACCATGAGAGAGCTG 365
Db 469 -----ValGly 470
QY 366 GCGCGAGAGCGCTGCGCGGACGATGCGGCGGCTTCCGCGAGGCGGCTTGAAG 425
Db 470 yProGlyGlyAlaThr--ProGly-----GlyGlyGlyAlaArgGlyGlySerG 487
QY 426 GAATTCATGTGTGCGGACCACTACTGATATCGCATGCTGACGCGGACGCGGCGCA 485
Db 487 LysGlyGlnGlyAlaAlaGlnGlyAlaAsp--ProAlaSerGlyGlyProProProAla 506
QY 486 CT-----GAGCTTCGATCAAGTCTGCTCTCAGAGAGCTTTC 524
Db 506 TgGlnProGlySerGlyGlyGlyAlaProLeuAsnHisLeu-----P 520
QY 525 CTGTTGATGTGATCAAGCAGATGCTTCTGCGCGGTGTTTGGCTCAGGCTCCGACGG 584
Db 520 roThrGlnAl-----AlaGlyArg-HisProAlaThrProProGly 534
QY 585 CAGAGGCGCTAAGCCCAAGCTTGGCGGCTTCC 617
Db 535 AlaHisGlyHisAlaGlnProProAlaProGlyTyAspProAlaProAlaThrHisAla 554
QY 617 ----- 617
Db 555 SerGlnProGlyHisGlnHisProTyAspGlyGlyGlyLeuGlyProThrProPro 574
QY 618 TAGGTCATGCTCTCTCCCTAGGG-----AATGTCACGACAG 656
Db 575 GlnGlyProProProProTyGlyProProHisGlnProValArgAsnGlyArgSerThr 594
QY 657 AGTGGCAGATTCATGTGGGCGGCGCTGATGTTGCTGCGAGAGAGGACGCTTACATG 716
Db 595 AlaLeuLeuValValAlaLeuValAlaLeuGlyAlaGlySerValTyAla 614
QY 717 TTG 719
Db 615 Leu 615

```

RESULT 11

S24407 Formin isoform IV - mouse

C/Species: Mus musculus (house mouse)

C/Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999

C/Accession: S24407

R/Jackson-Grusby, L.; Xu, A.; Leder, P.

Genes Dev. 6, 29-37, 1992

A/Title: A variant limb deformity transcript expressed in the embryonic mouse limb def;

A/Reference number: S24407; MIMD:92112033; PMID:1339380

A/Molecule type: mRNA

A/Residues: 1-1206 <JAC>

A;Title: The covalent structure of collagen: amino acid sequence of alpha1-CB3 from calf
 A;Reference number: A91387; PMID:73049499; PMID:4673951
 A;Accession: A91387
 A;Molecule type: protein
 A;Residues: 146-294 <F12>
 A;Experimental source: skin
 R;Fletterick, P.P.; Rexrodt, F.W.; Hopper, K.E.; Kuehn, K.
 Eur. J. Biochem. 38, 396-400, 1973
 A;Title: The covalent structure of collagen. 2. The amino-acid sequence of alpha1-CB7 from calf
 A;Reference number: A91211; PMID:74086118; PMID:4359390
 A;Accession: A91211
 A;Molecule type: protein
 A;Residues: 295-562 <F13>
 A;Experimental source: skin
 R;Wendt, P.; Mark, K.V.D.; Rexrodt, F.; Kuehn, K.
 Eur. J. Biochem. 30, 169-183, 1972
 A;Title: The covalent structure of collagen. The amino-acid sequence of the 112 residues
 A;Reference number: A91201; PMID:73042276; PMID:4343808
 A;Accession: A91201
 A;Molecule type: protein
 A;Residues: 563-675 <MEN>
 A;Experimental source: skin
 R;Fletterick, P.P.; Rexrodt, F.W.; Wendt, P.; Stark, M.; Kuehn, K.
 Eur. J. Biochem. 30, 163-168, 1972
 A;Title: The covalent structure of collagen. Amino acid sequence of peptide alpha1-CB6 from calf
 A;Reference number: A91200; PMID:73042275; PMID:4343807
 A;Accession: A91200
 A;Molecule type: protein
 A;Residues: 676-758 <F14>
 A;Experimental source: skin
 R;Fletterick, P.P.; Rexrodt, F.; Fletterick, J.; Fletterick, P.; Stark, M.; Kuehn, K.
 FEBS Lett. 21, 75-79, 1972
 A;Title: The amino acid sequence of the carboxyterminal nonhelical cross link region of alpha1(CB3) from calf
 A;Reference number: A43048
 A;Accession: A43048
 A;Molecule type: protein
 A;Residues: 759-779 <RA2>
 A;Experimental source: skin
 C;Comment: Lysines at positions 115, 124, 274, 346, 424, 496, 658, and 670 may be hydroxylated
 C;Comment: The order of the eight CNBR peptides in the alpha 1(I) chain of bovine skin is 9, 149, 268, and 217 residues.
 C;Comment: The complete chain contains 1052 residues.
 C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology; C;Keywords: coiled coil; extracellular matrix; glycoprotein; pyroglutamic acid; trimer; F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Alignment Scores:
 Pred. No.: 1.04e-05 Length: 779
 Score: 184.00 Matches: 87
 Percent Similarity: 36.55% Conservative: 19
 Best Local Similarity: 30.00% Mismatches: 94
 Query Match: 13.16% Indels: 90
 DB: 1 Gaps: 13

US-10-023-182-1 (1-752) x CGB01S (1-779)

QY 683 TCAGCCCCCAATGACATGCGCCTGCTGGACCATCTCCAGGAGGAGGAGCAT 624
 Db 214 AlAGlyProProGlyAla-----GlyLysProGly 224
 QY 623 GACCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 564
 Db 225 GluGlnGlyValProGlyAspLeuGly--AlaProGlyProSerGlyAlaValGlyGln 243
 QY 563 AAA-----CACGGGAGAAAGCACTGCTGATCCATCAAGGAGGAGGAGGAG 510
 Db 244 ArgGlyPheProGlyGluArgGlyValGluGlyPro-----ProGlyProAlaGlyPro 261
 QY 509 ACAGAGCTATGAGAGGCTGACGCTGCGGCTGCTGCGAGCACTAGTGGATGATGAG 450
 Db 262 ArgGlyAlaGlnGly--AlaProGlyAsnAspGlyAlaValGlyAspAlaGly----- 278

QY 449 TATGTTGCCGAGACAGAGTGAATCTCTTCAAGAGCACTCTGGGACGGGAAACGGTGGCC 390
 Db 279 -----AlaProGlyAlaProGlySerGlnGly 288
 QY 389 ATCCGAGG-----CAAGGCTCTGCGGGAGGCTGCTGCTCAATGAGGTGCGC 342
 Db 288 lAProGlyLeuGlnGlyMeProGlyGluArgGlyAlaAlaGlyLeuPro-GlyProLys 307
 QY 341 GAAAGCATGCGGAGGTGAATCTCAAGAGGCG----- 309
 Db 308 GlyAspArgGlyAspAlaGlyProLysGlyAlaAspGlyAlaProGlyLysAspGlyVal 327
 QY 308 -----GCTTCGCGGCGCCCTGGCGCCGACCTGAGCATCATTCAGCC 264
 Db 328 ArgGlyLeuThrGlyProLysGlyProGlyProGlyProAlaGlyAlaProGlyAspLys-Gl 347
 QY 263 TGAAGCCGCGCG-----CCATGCGGACCCCGCGGCGCGCT----- 227
 Db 347 YGluAlaGlyProSerGlyProAlaGlyThrArgGlyAlaProGlyAspArgGlyLys 367
 QY 226 -----CTTCGCGGCGCCGAGGCGCTTGTGCGCCCTGCGCGGAGACTGCGCGCGCT 171
 Db 367 OGlyProProGlyProAlaGlyPheAlaGlyPro--ProGlyAlaAspGlyGlnProG 386
 QY 170 GGACCGCGCTCTCTGCGCGCG-----CCAGCATTTGCCCTTGGGCGCATC 126
 Db 386 YAlaLysGlyGluProGlyAspAlaGlyAlaLysGlyAspAlaGlyProProGlyProAl 406
 QY 125 AGAATGCCAGGCGCTCTGCGCGCA----- 101
 Db 406 AGlyProAlaGlyProProGlyProLysGlyAsnValGlyAlaProGlyProLysGlyAl 426
 QY 100 -----TCAGCATCGCGCC-----GTCGA 84
 Db 426 ArgGlySerAlaGlyProProGlyAlaThrGlyPheProGlyAlaAlaGlyArgValG 446
 QY 83 ACCCGCTGCGCGCGCGCTTGGCGCTGATGCTGCGAGCGCTTGGCGCTGCGAG 24
 Db 446 YProProGlyProSerGlyAsnAlaGlyProProGlyProProGlyProAlaGlyLysG 466
 QY 23 AGAAGTACGCGCGCGCGAGGA 2
 Db 466 uGlySerLysGlyProArgGly 473

RESULT 9
 T15936
 Hypothesized protein EGAP7.1 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000
 C;Accession: T15936
 R;Miller, N.
 submitted to the EMBL Data Library, May 1996
 A;Description: The sequence of C. elegans cosmid EGAP7.
 A;Reference number: Z18433
 A;Accession: T15936
 A;Status: preliminary; translated from GB/EMBL/DBU
 A;Molecule type: DNA
 A;Residues: 1-302 <M11>
 A;Cross-references: EMBL:U58736; NID:G1326297; PID:G1326298; PIDN:AA00598.1; GSPDB:GN01
 C;Experimental source: strain Bristol N2; clone EGAP7
 C;Genetics:
 A;Gene: CESP:EGAP7.1
 A;Map position: X
 A;Introns: 55/3
 C;Superfamily: unassigned collagens

Alignment Scores:
 Pred. No.: 1.23e-05 Length: 302
 Score: 183.50 Matches: 62
 Percent Similarity: 36.46% Conservative: 8
 Best Local Similarity: 32.29% Mismatches: 63
 Query Match: 13.13% Indels: 59

A>Note: the authors translated the codons CAG for 721 and CGT for 738 as Glu
 R:Polino, A.; Zolzezi, F.; Valli, M.; Pignatti, P.F.; Cetta, G.; Brunelli, P.C.; Montes
 Hum. Mol. Genet. 3, 2201-2206, 1994
 A:Title: Severe (type III) osteogenesis imperfecta due to glycine substitutions in the C
 A:Reference number: 154365; PMID:95187161; PMID:7881420
 A:Accession: 154365
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 746-766, 'S', 768-781 <FOR>
 A:Cross-references: GB:L47667; NID:G1009093; PIDN:AAB59576.1; PID:G1009094
 A:Chesler, S.D.; Wallis, G.A.; Byers, P.H.
 J. Biol. Chem. 268, 18218-18225, 1993
 A:Title: Mutations in the carboxyl-terminal propeptide of the pro alpha 1(I) chain of ty
 A:Reference number: A47426; PMID:93352646; PMID:8349697
 A:Accession: A47426
 A:Molecule type: mRNA
 A:Residues: 1179-1276, 'H', 1278-1336, 1339-1387, 'R', 1389-1464 <CHE>
 A:Cross-references: GB:S64596; NID:G407589; PIDN:AAB27856.1; PID:G407590
 A:Note: sequence extracted from NCBI backbone (NCBI:136444, NCBI:P136445)
 A:Note: does not represent an experimentally determined sequence but three different mut
 A:Accession: B47426
 A:Molecule type: mRNA
 A:Residues: 1179-1464 <CH4>
 A:Experimental source: normal dermal fibroblast culture
 A:Accession: C47426
 A:Molecule type: mRNA
 A:Residues: 1179-1276, 'H', 1278-1464 <CH5>
 A:Experimental source: fetal cell 86-237
 A:Accession: D47426
 A:Molecule type: mRNA
 A:Residues: 1179-1336, 1339-1464 <CH6>
 A:Experimental source: fetal cell 86-146
 A:Accession: E47426
 A:Molecule type: mRNA
 A:Residues: 1179-1387, 'R', 1389-1464 <CH7>
 A:Experimental source: fetal cell 88-251
 R:Cohn, D.H.; Apone, S.; Eyre, D.R.; Starman, B.J.; Andressen, P.; Charbonneau, H.; Nid
 J. Biol. Chem. 263, 14605-14607, 1988
 A:Title: Substitution of cysteine for glycine within the carboxyl-terminal Telopeptide c
 A:Reference number: 155269; PMID:89008319; PMID:3170557
 A:Accession: 155269
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1187-1194, 'C', 1196-1220 <COH>
 A:Cross-references: GB:M23213; NID:G340842; PIDN:AAB59363.1; PID:G499622
 A:Note: mutant sequence from a patient with mild osteogenesis imperfecta
 R:Maekela, U.K.; Kaasina, M.; Virts, A.; Vuorio, E.
 Nucleic Acids Res. 16, 349, 1988
 A:Title: Human pro-alpha-1(I) collagen: cDNA sequence for the C-propeptide domain.

Alignment Scores:

Pred. No.:	7,63e-06	Length:	1464
Score:	185.50	Matches:	87
Percent Similarity:	36.868	Conservative:	21
Best Local Similarity:	29.698	Mismatches:	92
Query Match:	13.27%	Indels:	93
DB:	1	Gaps:	13

US-10-023-182-1 (1-752) x CGHUIS (1-1464)

```

QY 683 TCAGGCCCCCAATGATGCGCACTGCTGCGGACCATTCCTTAGGGAGAGAGCAT 624
DB 649 AAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
QY 623 GACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 660 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 563 AAA-----CAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 679 ATGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
QY 509 ACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

```

```

DB 697 ArgGlyAlaLeuGly--AlaProGlyAsnAspGlyAlaLeuGlyAspAlaGly----- 713
QY 449 TATGTGGCGGACACAGTGAATCTCTTACAGAGACCCCTGACAGGAGAGAGAGAGAGC 390
DB 714 -----AlaProGlyAlaProGlySerLeuGly 723
QY 389 ATCTGTGG-----CCAGCTCTCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 342
DB 723 LaProGlyLeuGlnGlyMetProGlyLeuArgGlyAlaAlaGlyLeuPro-GlyProlys 742
QY 341 GAAAGCATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 309
DB 743 GlyAspArgGlyAspAlaGlyProGlySerLeuAlaAspGlySerProGlyLeuAspGlyVal 762
QY 308 -----GCTCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 264
DB 763 ArgGlyLeuThrGlyProGlyLeuGlyProGlyProGlyProGlyAlaProGlyAspLys-Gl 782
QY 263 TGAAGCCGCGCG-----CCATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 227
DB 782 YGluSerGlyProSerGlyProAlaGlyProThrGlyAlaArgGlyAlaProGlyAspPar 802
QY 226 -----CCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
DB 802 GgylGluProGlyProGlyProGlyProGlyProGlyProGlyProGlyProGlyProGlyPro 821
QY 179 GCGGCGCGTGGACACCGGCTCTCTGCGCG-----CCAGCATTCGCGCG 135
DB 821 YGluProGlyAlaLeuGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 841
QY 134 TGGGCGCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 101
DB 841 GgylProAlaGlyProAlaGlyProProGlyProGlyProGlyProGlyProGlyProGlyPro 861
QY 100 -----TCGATCGCGC----- 89
DB 861 AluGlyAlaArgGlySerAlaGlyProProGlyAlaThrGlyPheProGlyAlaAlaGly 881
QY 88 -----GTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 33
DB 881 YArgValGlyProProGlyProSerGlyAsnAlaGlyProProGlyProProGlyProAl 901
QY 32 CTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2
DB 901 AglyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 911

```

RESULT 8

CGS01S
 collagen alpha 1(I) chain - bovine (tentative sequence) (fragments)
 C/Species: Bos primigenius taurus (cattle)
 C/Date: 24-Apr-1984 #sequence revision 31-Dec-1993 #ext change 31-Mar-2000
 C/Accession: A91193; A91229; A91387; A91211; A91201; A91200; A43048; A02853
 R:Rautenberg, J.; Timpl, R.; Furthmayr, H.
 Eur. J. Biochem. 27, 231-237, 1972
 A:Title: Structural characterization of N-terminal antigenic determinants in calf and h
 A:Reference number: A91193; PMID:7225354; PMID:4115172
 A:Accession: A91193
 A:Molecule type: protein
 A:Residues: 1-19 <RAU>
 A:Experimental source: skin
 A:Note: the epsilon carbon of Lys-9, by homology with the rat alpha 1(I) chain, is conv
 R:Rietzek, P.P.; Kuehn, K.
 Eur. J. Biochem. 52, 77-82, 1975
 A:Title: The covalent structure of collagen: amino-acid sequence of the cyanogen-bromid
 A:Reference number: A91229; PMID:7602320; PMID:1164916
 A:Accession: A91229
 A:Molecule type: Protein
 A:Residues: 20-145 <PIE>
 A:Experimental source: skin
 A:Note: Lys-103 is hydroxylated and binds glucosylgalactose
 R:Rietzek, P.P.; Wendt, P.; Kuehn, K.
 FEBS Lett. 26, 74-76, 1972

A/Reference number: 160114; MUID:86329734; PMID:2843432
A/Accession: 160114
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-369, 'L', 371-589 <DAL>
A/Cross-references: GB:M20789; NID:g179593; PIDN:AAB59373.1; PID:g179594
A/Note: Submitted to the EMBL/GenBank/DBJ databases by Prockop, D.J., 13-JUN-1988
A/Title: Human proalpha1(I) collagen gene structure reveals evolutionary conservation of
A/Reference number: 501143; MUID:89025644; PMID:3178743
A/Accession: 501143
A/Molecule type: mRNA
A/Residues: 1-472 <TRC>
A/Cross-references: EMBL:X07884; NID:g30015; PIDN:CA30731.1; PID:g30016; GB:M36546; NID
A/Note: Submitted to the EMBL/GenBank/DBJ databases by Prockop, D.J., 13-JUN-1988
A/Title: Human proalpha1(I) collagen gene structure reveals evolutionary conservation of
A/Reference number: A93335; MUID:84270697; PMID:6462220
A/Accession: A93335
A/Molecule type: DNA
A/Residues: 1-58, 'Q', 60-181 <CHU>
A/Cross-references: EMBL:X00820; NID:g35657; PIDN:CAA25394.1; PID:g35658
A/Note: Submitted to the EMBL/GenBank/DBJ databases by Prockop, D.J., 13-JUN-1988
A/Title: DNA sequences in the first intron of the human pro-alpha 1(I) collagen gene en
A/Reference number: 155254; MUID:88033098; PMID:2822714
A/Accession: 155254
A/Status: translation not shown; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-45 <ROS>
A/Cross-references: GB:U02829; NID:g180387; PIDN:AA5193.1; PID:g180388
A/Note: Submitted to the EMBL/GenBank/DBJ databases by Prockop, D.J., 13-JUN-1988
A/Title: Regulatory elements in the first intron contribute to transcriptional control
A/Reference number: A39943; MUID:88037389; PMID:3480516
A/Accession: A39943
A/Molecule type: DNA
A/Residues: 1-34 <BOB>
A/Cross-references: GB:U03559; NID:g180876; PIDN:AA52052.1; PID:g553238
A/Note: Submitted to the EMBL/GenBank/DBJ databases by Prockop, D.J., 13-JUN-1988
A/Title: Fine structural analysis of the human pro-alpha 1 (I) collagen gene. Promoter
A/Reference number: 155237; MUID:85130970; PMID:2857713
A/Accession: 155237
A/Status: translation not shown; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-34 <CH2>
A/Cross-references: GB:M10627; NID:g180383; PIDN:AA5192.1; PID:g55326
A/Note: Submitted to the EMBL/GenBank/DBJ databases by Prockop, D.J., 13-JUN-1988
A/Title: A base substitution in the exon of a collagen gene causes alternative splicing
A/Reference number: S09400; MUID:89356643; PMID:2767050
A/Accession: S09400
A/Molecule type: mRNA
A/Residues: 156-183 <WET>
A/Note: Submitted to the EMBL/GenBank/DBJ databases by Prockop, D.J., 13-JUN-1988
A/Title: Isolation and characterization of the cyanogen bromide peptides from the alpha
A/Contents: CNBR0-1, CNBR2, CNBR4, CNBR5
A/Accession: B90567
A/Molecule type: protein
A/Residues: 162-198, 'Z', 200-201, 'Z', 203-206, 'Z', 208-209, 'Z', 211-228, 'B', 230, 'BB', 233, 'Z'

A/Experimental source: skin
A/Note: evidence for 170-alanine
A/Accession: B90567
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-45 <ROS>
A/Cross-references: GB:U02829; NID:g180387; PIDN:AA5193.1; PID:g180388
A/Note: Submitted to the EMBL/GenBank/DBJ databases by Prockop, D.J., 13-JUN-1988
A/Title: Regulatory elements in the first intron contribute to transcriptional control
A/Reference number: A39943; MUID:88037389; PMID:3480516
A/Accession: A39943
A/Molecule type: DNA
A/Residues: 1-34 <BOB>
A/Cross-references: GB:U03559; NID:g180876; PIDN:AA52052.1; PID:g553238
A/Note: Submitted to the EMBL/GenBank/DBJ databases by Prockop, D.J., 13-JUN-1988
A/Title: Fine structural analysis of the human pro-alpha 1 (I) collagen gene. Promoter
A/Reference number: 155237; MUID:85130970; PMID:2857713
A/Accession: 155237
A/Status: translation not shown; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-34 <CH2>
A/Cross-references: GB:M10627; NID:g180383; PIDN:AA5192.1; PID:g55326
A/Note: Submitted to the EMBL/GenBank/DBJ databases by Prockop, D.J., 13-JUN-1988
A/Title: A base substitution in the exon of a collagen gene causes alternative splicing
A/Reference number: S09400; MUID:89356643; PMID:2767050
A/Accession: S09400
A/Molecule type: mRNA
A/Residues: 156-183 <WET>
A/Note: Submitted to the EMBL/GenBank/DBJ databases by Prockop, D.J., 13-JUN-1988
A/Title: Isolation and characterization of the cyanogen bromide peptides from the alpha
A/Contents: CNBR0-1, CNBR2, CNBR4, CNBR5
A/Accession: B90567
A/Molecule type: protein
A/Residues: 162-198, 'Z', 200-201, 'Z', 203-206, 'Z', 208-209, 'Z', 211-228, 'B', 230, 'BB', 233, 'Z'

QY	6	CGGGGGCCCTGACCTTCTCTCGAAGAGCCGGGCGAGAGCTCCGAGCGACGAGGCGGAA	65
Db	987	ArgGlyProAlaPro-----ProArgGlyArgGlyAlaGlyArgAlaLeu	1002
QY	66	GGCGGGGGCAKAGGGGGTTG-----ACGGGCGCATGCTGATGCG	104
Db	1003	PrometGlyGlyGlyAspGluAlaProLeuProValProThrGlyProSerGlyGly	1022
QY	105	CCAGAG-----GGCCCTTGACATTCGTATGCGCCAGGGGGCAATGCTGAGCGC	152
Db	1023	ProGlyProAlaGlyProGlyArgGlyMetProAlaGlyProGlyArgGlyAlaProGly	1044
QY	153	CCAGAGAGAGCGGGTGCCACCGGGCGGACAGAGTCCCGGGGGCCGACGGGGCAGCAGGCGC	212
Db	1043	ProGlyArgGlyGlyProGlyGlyProProGlyProAlaGlyProGlyArgGly--Ala	1061
QY	213	TCGGGGCCGAGAGAGCGCGCCG-----CGGGGTCCGACATGCG-----	251
Db	1062	ProGlyProGlyArgGlyAlaProGlyProSerArgGlyGlyProGlyGlyProProPro	1083
QY	252	-----GGCGGGCTTCAGGGCTGAAATGATGCTGCAGATGC	287
Db	1082	GlyGlyArgGlyMetProProProProGlyAlaArgGlyGlyProGlyGly--ProGlyProAla	1101
QY	288	GGGGCGCAGGGGCGGAGAGCGCGCTGTTAGTTCACTCGCATGCCCTTTCGGACA	347
Db	1101	agGlyProGlyArgGlyMet--ProAla-----P	1111
QY	348	CCCATGAAACAGAGCTGGCCCGCAGCAGCTGGGCCA-----	385
Db	1111	roGlyGlyArgGlyArgGlyGlyProGlyGlyProAlaGlyArgGlyGlyProGlyP	1133
QY	386	-----GGATGCCCAACCGCTTCCCGTCGACAGGGGTGCTTGAAGAGTTCACT	434
Db	1131	roAlaGlyArgGlyMetProProProGlyAlaGlyArgGlyGly--	1145
QY	435	GTCGCGGCAATCTACTATCCGACTGACTCTGCACAGCAGCGCCACTGCAGCT	493
Db	1146	-----ProAlaProAlaAlaAla	1151

RESULT 6

T22602

hypothetical protein F54B11.2 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T22602

R:Swingburne, J

submitted to the EMBL Data Library, March 1996

A:Reference number: Z19588

A:Accession: T22602

A:Structure: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-304 <EIL>

A:Cross-references: EMBL:Z70208; PIDN:CAA4316.1; GSPDB:GN00028; CESP:F54B11.2

A:Experimental source: clone F54B11

C:Genetics:

/gene: CESP:F54B11.2

QY 528 ACAGGAAAGTGTCTGGAGACAGAGCTGATGAGAGCTGACATTGGCGGTGCTGGAG 466

Db 78 ThrAlaArgGlnIaGlyTYrThrSer-----GlyGlyGlyAlaAsp 91

QY 468 CAGTCAGTCGGATATGTCAGTA-----TGTGCCGACACAGTAAGT 427

Db 92 AlaGlyAlaGlyAlaSerSerGlyGlyGlnCysGlyCysAsn----- 107

QY 426 CTTTCAGAGAGCACCCTTGACACAGGAACCGGTGGCATCTCTGGG---CGAGCTCTGC 370

Db 108 -----ProGlyProProGlyYalAlaGlyAsnProGlyTyLysProGlyTyLys--- 122

QY 369 GGGCAGACTCTGCTTCATGCTGCTCCGGAAGAGCATGCGAGTAGAACCTCAGACAGG 310

Db 123 -----ProGlyTyLysProGlyYalAlaPro 129

QY 309 GGCCTCTCG---GCCCTCTGCCCCGATCTGCAGATCATTCAGCCCTGAAGCGGAGC 253

Db 130 GlyAsnProGlyYalAla-ProGlyTyLysGlyAlaAlaVal-ProCysGlnAlaTyThrProP 149

QY 252 CGGCATCGAGACCCCGGGGGCGCTCTCCCGGCGCCGAG----- 212

Db 149 roProGlyLysProCysProAlaGlyProProGlyProProGlyProAlaAspGlyProAlaG 169

QY 211 -----GCCCTTCTGCTGCCCTCTGCGCCCGGAGGACTC 181

Db 169 lyProAlaGlyProAlaAspGlyGlnAlaGlySerProAlaAlaProSerProProGlyProP 189

QY 180 TGCGCGCCCGTGGACCGGCC-----TCTCTCTGGAG 151

Db 189 roGlyProSerGlyProAlaGlyProAlaGlyAsnAspGlyYalAlaGlyThrProGlyP 209

QY 150 CG-----CCAGCATTGC 139

Db 209 roAspGlyProAlaGlyGlnSerThrTYrProGlnProAlaAlaProGlyProAlaGlyP 229

QY 138 CCGCTGGGCGCATCGAGATGCGAGGCTCTCTGGG-----CCATCAGCATGCGCC--- 89

Db 229 roProGlyProAlaGlyProProGlyProAlaAspGlyYalAlaSerProThrAlaAlaProGlyY 249

QY 88 --GTGGAACCCCT--GTGCCCGGCTTGCGCTGCTGATGCTCGAGACCTTGCCCG 34

Db 249 laAlaGlyProProGlyProProGlyProAlaGlyYalAlaProGlyProAlaGlyAsnProG 269

QY 33 GCTCT-----CAGAGAGAGGTCAGGCGCCACGAGGA 2

Db 269 lyThrAlaGlyProProGlyYalAlaGlyGlyGlnGlyGlnTyLysGly 284

RESULT 7

CGH015

collagen alpha 1(I) chain precursor - human

N/Alternate names: procollagen alpha 1(I) chain

C/Species: Homo sapiens (man)

C/Date: 12-Aug-1981 #sequence revision 04-Oct-1996 #text change 31-Dec-2000

ClAccession: I50114; S01143; A93335; I55224; A39943; I55237; S09400; B90567; S1

5269; A29439; I53466; A08852; I37247

R/D/Accessio M.; Bernard, M.; Pretorius, P.J.; de Wet, W.; Ramirez, F.; Pretorius, P.J.

Gene 67, 105-115, 1988

Title: Complete nucleotide sequence of the region encompassing the first twenty-five

R.Fenton, S.P.; Lemande, S.R.; Hannagan, M.; Stacey, A.; Jaenisch, R.; Bateman, J.F.
 Biochim. Biophys. Acta 1216, 469-474, 1993
 A:Title: Genomic sequence of mouse COL1A1 encoding the collagen propeptides.
 A:Reference number: 539789; MUID:94092741; PMID:8268229

A:Accession: S39789
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-185, 1031-1201, 'G', 1203-1218, 'E', 1220-1221, 'T', 1223-1
 R/Rhodes, K.; Ripe, R.A.; Umezawa, A.; Nehls, M.; Brenner, D.A.; Breindl, M.
 Mol. Cell. Biol. 14, 5950-5960, 1994

A:Title: DNA methylation represses the murine alpha 1(I) collagen promoter by an indirect
 A:Reference number: 148300; MUID:94344105; PMID:8065328

A:Accession: 148300
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>

A:Cross-references: EMBL:X54876; NID:950486; PIDN:CA38657.1; PID:950487
 C:Genetics:

A:Gene: COL1A1

A:Intons: 770/3, 786/3, 806/3, 842/3, 860/3, 878/3, 932/3, 968/3, 1004/3, 1022/3, 1058/
 C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; heterotrimer; triple helix
 F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-151/Domain: amino-terminal propeptide #status predicted <PRO>
 F:30-89/Domain: von Willebrand factor type C repeat homology <VWC>
 F:152-1453/Product: collagen alpha 1(I) chain #status predicted <MAT>

F:1224-1453/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Alignment Scores:

Pred. No.: 5,05e-06 Length: 1453
 Score: 188.00 Matches: 66
 Percent Similarity: 37.31% Conservative: 9
 Best Local Similarity: 32.84% Mismatches: 62
 Query Match: 13.45% Indels: 64
 Gaps: 8

US-10-023-182-1 (1-752) x S21626 (1-1453)

QY 417 GCACCCCTGGACGCGGAGGCGATCTCTGCG-----CGAGCTCTCG 370
 Db 703 AAlProGlyAlaProGlySerGlnGlyAlaProGlyLeuGlnGlyMetProGlyGluArg 722
 QY 369 GGGCAGCTCTGCTTCATGAGGTGTGCGAAGGCGATGGAGGTAGAACTCAAGAGCG 310
 Db 723 G1YAlaAlaGlyLeuPro-GlyProGlyGlyAspArgGlyAspAlaGlyProGlyGlyAl 742
 QY 309 G-----GCTCTCCGCGCCCTCG 292
 Db 742 AAspGlySerProGlyLysAspGlyAlaArgGlyLeuThrGlyProGlyLeuProGly 762
 QY 291 CCCGCGATCTGCGAGCATCTTCAAGCCCTGAGCCGCGCG-----CAT 247
 Db 762 YProAlaGlyAlaProGlyAspLys-GlyGluAlaGlyProSerGlyProProGlyPro 782
 QY 246 GCGGACCCCGCGGCGCGCT-----CCTCCGCGCGCCGAGGCC 208
 Db 782 hrg1yAlaAlaGlyAlaProGlyAspArgGlyGluAlaGlyProProGlyProAlaGly 802
 QY 207 TTGCTGCCCT----- 197
 Db 802 hnalaglyProProGlyAlaAspGlyGlnProGlyAlaLysGlyGluProGlyAspThr 822
 QY 196 -----GCGCCCGCGGAGCTCTGCGCGCGCGCGAGCCGCTCTCT 155
 Db 822 lyVallysglyAspAlaGlyProProGlyProAlaGlyProAlaGlyProProGlyPro 842
 QY 154 -----GGCGCGCGAGCATTTGCGCGCGCGCGCGAGATGCGAGGCGCTCTG 106
 Db 842 leglyAsnValGlyAlaProGly---ProGlyGlyProAlaGlyAlaAlaGlyProPro 861
 QY 105 GGGCATCAGATGCGCC-----GTGACCCCTGCGCCCGCGCTTGG 61
 Db 861 lyAlaThrGlyPheProGlyAlaAlaGlyArgValGlyProProGlyProSerGlyAsn 881

QY 60 CCTGCATGAGCTCCGAGAGCTGCGCGCTCGCGAGAGAGAGGCGAGGCGCCAGAGGA 2
 Db 881 lAGlyProProGlyProProGlyProValGlyLysGlyGlyGlyGlyProAlaGly 900

RESULT 4

T35608
 polypeptide hydroxylase - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 17-Mar-2000

C:Accession: T35608
 R:Seeger, K.O.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, June 1999

A:Reference number: 221584
 A:Accession: T35608

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 1-627 <SER>

A:Cross-references: EMBL:AL079356; PIDN:CA45603.1; GSPDB:GN00070; SCOEDB:SC669.12C
 A:Experimental source: strain A3(2)
 C:Genetics:

A:Gene: SCOEDB:SC669.12C
 C:Superfamily: tetracycline 6-hydroxylase

Alignment Scores:

Pred. No.: 6,99e-06 Length: 627
 Score: 186.50 Matches: 50
 Percent Similarity: 45.83% Conservative: 5
 Best Local Similarity: 41.67% Mismatches: 34
 Query Match: 13.25% Indels: 31
 Gaps: 6

US-10-023-182-1 (1-752) x T35608 (1-627)

QY 36 GCGAGAGCTCCGAGCGCATCGAGCCGAGCGCGGCGAGCGGCGTTGACGCGCAT 95
 Db 399 GlyProGlyAlaGlyThrProGlyGlyAlaGlyArgGlyThrGlyGly---ProGlyGly 417

QY 96 GGTATGCGCCAGAGAGCGCTCGCATTCATGCGCCAGAGGCGCATGCTGCGGCCA 155
 Db 418 ProGlyGlyProGlyGlyLeuGlyGlyProGlyGlyProGlyGly---ThrGlyGlyPro 436

QY 156 GAGAGGCGGCTGCGACCGCGCGC---AGAGTCCCGCGGCGCGAGGCGAGAGGCC 212
 Db 427 GlyGlyProGlyGlyProGlyGlyProAspGlyProArgGlyAlaGlyAlaProGly 456

QY 213 TCGGCGCGGAGAGAGCGCGCGCGCGGTCCGATGCGCGCGCGCTTCAAGGCTGAT 272
 Db 457 GlyGlyProGlyGly---GlyProGlyGlyGly----- 466

QY 273 GGATCTGAGATGCGGCGCGAGGCGCGAGCGCGCTGCTGAGTTCTACCTCGCC 332
 Db 467 -----GlyProGlyArgGlyLeuAsnValAlaLeuGly 478

QY 333 ATGCTTTC-----GGACACCCAGAGAGAGAG 362
 Db 479 TyrArgTyrProArgGlyAlaValAlaGlyAlaAspProAlaThrProValAlaProGlu 498

RESULT 5

T32734
 myosin-IIA - Acanthamoeba castellanii

C:Species: Acanthamoeba castellanii
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 08-Sep-2000

C:Accession: T32734
 R:Lee, W.L.; Ostap, E.M.; Zot, H.G.; Pollard, T.D.

submitted to the EMBL Data Library, August 1998

A:Reference number: 221216
 A:Accession: T32734

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 1-1215 <LEB>
 A:Cross-references: EMBL:AF085185; NID:93599477; PID:93599478; PIDN:AC35257.1

A:Experimental source: strain Neff

Db 102 -----ProGlyPro 104
 QY 372 TCGGGCCAGCTCTGCTTCATGGGTGTCGAAAGCATGCGAGTGAATCAAGCA 313
 Db 105 ProGlyProAlaGlyThrPro-----GlyLysProGlyLysProGlyLysAla 122
 QY 312 GCGGGCTCTCCGGC-----CCCTGGCCCGCATCTGCGAGCATTCATGAGCCCTGAAGCC 257
 Db 123 ProGlyLeuProGlyLysProGlyLysProPro---GlnGlnProGlyLysProThr 141
 QY 256 GCGCCGCCATGCGAAGCCCGCGAGGCGCTCTCCCGGCGCCGAGCCCTGCTGCTCC 197
 Db 142 ProProProGlyLysProGlyLysProGlyLysProPro-----Pro 156
 QY 196 GCGCCCGGAGCCTCTGCGCCCGCGAGGCGCTCTCCCGGCGCCGAGCATTCGCTCC 137
 Db 157 GlyProProGly-----ProProGlyLysSerGlyLysProGlySerProGlyLysPro 174
 QY 136 -----CCTGGGCCATGAGATGCGAGGCTCTGCTGGG--- 104
 Db 175 GlyLysAlaAlaAlaProGlyLysProGlyLysProGlyLysProProGlyLysAla 194
 QY 103 CCATCAGCA-----TCGCGCCCTCAAGCCCTCTGCGCCGCTTCCGCTGAGTGGCT 50
 Db 195 ProGlyAlaProGlyLysProGlyLysProGlyLysProAlaGlnSerGlnProLeu 214
 QY 49 CCG---GAGCTCTGCGCCGCTCTCTCAGAGAGAGTCAAGGCGCCAGAGCA 2
 Db 215 ProGlyGlnProGlyLysProGlyLysAlaGlyLysProGlyLysProProGly 231

RESULT 2

S57447
 HPRRII-7 protein - human
 N/Alternate names: HPRRII-4 protein
 C/Species: Homo sapiens (man)
 C/Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 24-Sep-1999
 C/Accession: S57447, S57489
 R/Fleischauer, K.L.
 submitted to the EMBL Data Library, June 1992
 A/Reference number: S57447
 A/Accession: S57447
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-551 <FL2>
 A/Cross-references: EMBL:X67337; NID:g871298; PIDN:CAA47752.1; PID:g871299
 C/Genetics: 221/3
 A/Intons: 221/3
 C/Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein F/82-151/Domain: ribonucleoprotein repeat homology <RRM4>

Alignment Scores:

Pred. No.: 3, 97e-06 Length: 551
 Score: 190.00 Matches: 59
 Percent Similarity: 39.87% Conservativity: 46
 Best Local Similarity: 37.34% Mismatches: 49
 Query Match: 13.59% Indels: 49
 DB: 2 Gaps: 7

US-10-023-182-1 (1-752) x S57447 (1-551)

QY 412 CCGGACGGAAGCGGTGCGATCTCGGCGAGGCTCTCTGCTGCTTC 353
 Db 208 ProGlyAlaValProGlyLysPargPhe-ProGlyProAlaGlyPro----- 223
 QY 352 ATGGGTGTCGGAAGCATGCGAGGTGAATCAAGAGCGGCTCTCCGCGCCCTG 293
 Db 224 -----GlyGlyProProPro-ProProP 230

QY 292 GCGCCGANTGTGAGCATTCATTAGCCCTGAAGCGCGCG-CCATCGGAGCC 239
 Db 230 heProAlaGly-GlnThrProProAlaProProLeuGlyProProGlyProProGlyPro 249
 QY 238 CCGGAGCGCGCTCTCTCCCGCGC-----CCGAGCGCCCTTGTCTGCGCCCGCG 188
 Db 250 ProGlyProProProProGlyLysValLeuProProProLeuAlaGlyProProAla 269
 QY 187 GGA-----CCTCTG-----CCGCGCGG 170
 Db 270 GlyAspArgProProProProValLeuPheProGlyLysProPheGlyLysProProLeu 289
 QY 169 GCACCGCGCTCTCTGCGCGCGCGAGATTGCCCTCGGCGCATCAAGATGCGAGGCT 110
 Db 290 GlyProLeuProProGlyProProProProProValProGlyLysGlyProProProGlyPro 309
 QY 109 CTTGGGCGCATGAGATCGCCCGCGAAGCCCTGTCGCG- 71
 Db 310 ProProProGlnGlnGlyProProProProProGlyProPheProProAlaProProGly 329
 QY 70 -----CGGCTTGGCGCATGCGCTCCGAGGCT---CTGCGCGCGC 32
 Db 330 ProLeuGlyProProLeuThrLeuAlaProProProProAlaProProGly 345

RESULT 3

S21626
 collagen alpha 1(I) chain precursor - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 13-Jan-1995 #sequence_revision 25-Apr-1997 #text_change 13-Aug-1999
 C/Accession: S57243, S16374, A23982, I49559, I49557, S39789, I48300, S21626
 R/Li, S.W.; Khillan, J.; Prockop, D.J.
 Matrix Biol. 14, 593-595, 1994
 A/Title: The complete cDNA coding sequence for the mouse pro-alpha-1(I) chain of type I
 A/Reference number: S57243
 A/Accession: S57243
 A/Molecule type: mRNA
 A/Residues: 1-1453 <LIS>
 A/Cross-references: EMBL:U08020; NID:g470673; PIDN:AAA86912.1; PID:g470674
 R/Metlaeranta, M.; Toivanen, D.; de Crombrughe, B.; Vuorio, E.
 Biochim. Biophys. Acta 1089, 241-243, 1991
 A/Title: Specific hybridization probes for mouse type I, II, III and IX collagen mRNAs.
 A/Reference number: S16176; MUID:91274355; PMID:2054384
 A/Accession: S16374
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1442-1453 <MET>
 A/Cross-references: EMBL:X57981; NID:g50484; PIDN:CAA41046.1; PID:g50485
 R/French, B.T.; Lee, W.H.; Maul, G.G.
 Gene 39, 311-312, 1985
 A/Title: Nucleotide sequence of a cDNA clone for mouse proalpha1(I) collagen protein.
 A/Reference number: A23982; MUID:86137403; PMID:3841523
 A/Accession: A23982
 A/Molecule type: mRNA
 A/Residues: 518-1128 <FRE>
 A/Cross-references: GB:M14423; NID:g192261; PIDN:AAA7333.1; PID:g192262
 R/Monson, J.M.; Friedman, J.; McCarthy, B.T.
 Mol. Cell. Biol. 2, 1362-1371, 1982
 A/Title: DNA sequence analysis of a mouse pro-alpha-1(I) procollagen gene: Evidence for
 A/Reference number: I49559; MUID:83141374; PMID:6298597
 A/Accession: I49559
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 735-1130 <RES>
 A/Cross-references: GB:M17491; NID:g192263; PIDN:AAA7333.1; PID:g192264
 R/Habbers, K.; Kuehn, M.; Delius, H.; Jaenisch, R.
 Proc. Natl. Acad. Sci. U.S.A. 81, 1504-1508, 1984
 A/Title: Insertion of retroviruses into the first intron of alpha1(I) collagen gene leads
 A/Reference number: I49557; MUID:84170331; PMID:6524198
 A/Accession: I49557
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-25 <R2>
 A/Cross-references: GB:K01688; NID:g192246; PIDN:AAA7330.1; PID:g553881

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 10, 2004, 17:32:16 ; Search time 21 seconds

(without alignments)
6887.509 Million cell updates/sec

Title: US-10-023-182-1
Perfect score: 1403
Sequence: 1 ATCCTCGTGGCCTGACCT.....TAAACTGAGTACGAAAA 752

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 566616

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-Q=/ggn2.1/USPRO.spool_P/US10023182/runat_10022004_171429_3891/app.query.fasta_1.903
-DB=FIR_76 -QFMT=fastan -SUPFIX=trp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR.SCOR=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10023182 -CGN_1_1_38@runat_10022004_171429_3891 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESUBSTRY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: PIR_76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	196	14.0	299	T29956	hypothetical prote
2	190	13.6	551	S57447	HBRII-7 protein
3	188	13.4	1453	S21626	collagen alpha 1(I
4	186.5	13.3	627	T35608	polyketide hydroxy
5	186.5	13.3	1215	T32734	myosin-1A - Acanth
6	186	13.3	304	T22602	hypothetical prote
7	185.5	13.3	1464	1 CGHUS	collagen alpha 1(I
8	184	13.2	779	1 CGBOIS	collagen alpha 1(I
9	183.5	13.1	302	2 T15936	hypothetical prote
10	183	13.0	783	2 T15389	hypothetical prote
11	183	13.1	1206	2 S24407	probable serine-th
12	183	13.1	1468	2 S11515	formin isoform IV
13	181	12.9	296	2 A31219	formin - mouse
14	179	12.8	1049	1 CGBO7S	collagen alpha 1(I

C 15	178.5	12.8	240	2 A24264	proline-rich prote
C 16	178.5	12.8	301	2 B31219	collagen 2 - Caeno
C 17	178.5	12.8	1532	2 A61262	collagen alpha 1(X
C 18	178	12.7	299	2 T24833	hypothetical prote
C 19	178	12.7	301	2 T21314	hypothetical prote
C 20	177	12.7	299	2 T19564	hypothetical prote
C 21	177	12.7	299	2 T23705	hypothetical prote
C 22	176.5	12.6	266	2 T23706	hypothetical prote
C 23	176.5	12.6	289	2 T25812	hypothetical prote
C 24	176.5	12.6	317	2 A28996	proline-rich prote
C 25	176	12.6	186	2 C41132	collagen-related p
C 26	176	12.6	299	2 T25407	hypothetical prote
C 27	175.5	12.6	227	2 C29149	proline-rich prote
C 28	175.5	12.6	240	2 B24264	proline-rich prote
C 29	175.5	12.6	278	2 B39066	proline-rich prote
C 30	175	12.5	316	2 T20497	hypothetical prote
C 31	175	12.5	473	2 S50755	hypothetical prote
C 32	174.5	12.5	298	2 UC1448	hypothetical prote
C 33	174.5	12.5	464	2 S22697	collagen col-34 -
C 34	174.5	12.5	641	1 QJBE31	nuclear antigen EB
C 35	174	12.4	587	2 T41653	extensin - Volvox
C 36	173.5	12.4	534	2 S21961	probable transcrip
C 37	173.5	12.4	1366	1 CGH2S	proline-rich prote
C 38	173.5	12.4	1880	2 T18531	collagen alpha 2(I
C 39	172.5	12.3	149	2 A41132	tracrin - medictin
C 40	171.5	12.3	1049	2 CGBO7S	collagen-related p
C 41	171.5	12.3	1446	2 A38587	collagen, cornea-s
C 42	171.5	12.3	1690	1 CGH1B	collagen alpha 4(I
C 43	171	12.2	675	2 S20819	collagen alpha 3(I
C 44	170.5	12.2	1464	2 S59856	collagen alpha 1(I
C 45	170	12.2	228	2 A44982	collagen UC011 - p

ALIGNMENTS

RESULT 1
T29956
hypothetical protein F36A4.10 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C/Accession: T29956
R/Miller, N.; Bradshaw, H.
Submitted to the EMBL Data Library, April 1996
A/Description: The sequence of C. elegans cosmid F36A4.
A/Reference number: Z20713
A/Accession: T29956
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-299 <MIL>
A/Cross-references: EMBL:U53333; PIDN:AAA96155.1; GSPDB:GN00022; CESP:F36A4.10
A/Experimental source: strain Bristol N2; clone F36A4
A/Genetics:
A/Gene: CESP:F36A4.10
A/Map position: 4
A/Intons: 27/3; 243/1
C/Superfamily: unassigned collagens

Alignment Scores:
Pred. No.: 1.56e-06
Score: 196.00
Percent Similarity: 40.68%
Best Local Similarity: 37.85%
Query Match: 14.02%
Gaps: 2
Indels: 43
Indels: 10

US-10-023-182-1 (1-752) x T29956 (1-299)
QY 492 GCTGCAGTTGGCGGTGCTGACGATGTCAGTATGTTCCGACACAG 433
DB 84 AaaalavalrhgylgylgylalaglnsergylserCysgluserCysleu----- 101
QY 432 TGAACCTCTTCAAGAACACCCCTGACGAGGAGGGGCAATCTGGGGCAAGCTCC 373

```

Db      142 GCGCGCTCTGCGCGAGGGGGTTCAGAAATGCGGCCCGCAATATTCACCTTCAGCGTGCCTT 201
Qy      340 TCGCACAACCCATGGAAGCAGAGCTGGCCCGCAGAGAGCTGGCCCGCAGATGCCCAACCGC 399
Db      202 TCCGACACCCCTTGAGAGGCGGAATCGCCCATGGGTCCCTGGACAGATGCCGAGCCGC 261
Qy      400 TTCCCGTGCAGAGGCTCTTGAAGAGTTCACGTGTCCGGAACAATGACTATCC 459
Db      262 ACCA-----AAGGtggttggaagatctcagatgtagggcagagatcctggtcgc 315
Qy      460 GACTGACTGCTGACAGACCAACCGCACTGACAGCTTCATGATGCTCTGTCCAGCAGC 519
Db      316 GGTGAAAGCTGAAGCTGTGCGCTGCTCCGAATTCGATCAATCTTCTTGACAGC 375
Qy      520 TTTCCTGTGATGTGATCAAGCATGCTTTCTG-CCCGTGTGTTGGCTCAGCTCC 578
Db      376 TTTCCTGTGTGTGCGACCATGACAGCGCTTGGGCCCGCGTTCCGCTAAGCTGGC 435
Qy      579 TCAGGCGACAGGCGCTAAGCCCGACGCTGGCGCCCTTCCTAGATGCTCTCCCTA 638
Db      436 CTGGGAATGAGAGAGGTCCCACTTGCCTCTCTGTAAGCAGTGGCCATCCTTC 495
Qy      639 GGGAAATG--GTCCAGACAGAGTGGCAGTTCAATGTGGGGGCTGATGTTGTGCTG 696
Db      496 CCTAGGCGAGAAATCCCAAGTGTCTACTTTCCTGGAGGGGCTCATGTTTATC---- 551
Qy      697 GAGGAGACGGCTTACATGTTGTTCTGTAGAAATAAATACTGAGCTA 745
Db      552 -----TGGTCTTAATGTTTGTCTTACTACAGAAATAAATACTGAGTA 594

```

RESULT 15

```

US-09-954-531-176
; Sequence 176, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 176
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-531-176

```

```

Query Match      15.6%; Score 117.4; DB 10; Length 599;
Best Local Similarity 59.5%; Pred. No. 4,9e-23;
Matches 279; Conservative 0; Mismatches 171; Indels 19; Gaps 4;

```

```

Qy      280 GAGATGCGGCGCGCGCGCGAGAGCGCGCTGTGAGTTTCACTTCGCGCAGCTT 339
Db      142 GCGCGCTCTGCGCGAGGGGGTTCAGAAATGCGGCCCGCAATATTCACCTTCAGCGTGCCTT 201
Qy      340 TCGCACAACCCATGGAAGCAGAGCTGGCCCGCAGAGAGCTGGCCCGCAGATGCCCAACCGC 399
Db      202 TCCGACACCCCTTGAGAGGCGGAATCGCCCATGGGTCCCTGGACAGATGCCGAGCCGC 261

```

```

Qy      400 TTCCCGTGCAGGCGGCTTCTGAAGAGTTCAGTGTGTCGCGCAACATGACTATGCC 459
Db      262 ACCA-----AAGGtggttggaagatctcagatgtagggcagagatcctggtcgc 315
Qy      460 GACTGACTGCTGACAGACCAACCGCACTGACAGCTTCATGATGCTCTGTCCAGCAGC 519
Db      316 GGTGAAAGCTGAAGCTGTGCGCTGCTCCGAATTCGATCAATCTTCTTGACAGC 375
Qy      520 TTTCCTGTGATGTGATCAAGCATGCTTTCTG-CCCGTGTGTTGGCTCAGCTCC 578
Db      376 TTTCCTGTGTGTGCGACCATGACAGCGCTTGGGCCCGCGTTCCGCTAAGCTGGC 435
Qy      579 TCAGGCGACAGGCGCTAAGCCCGACGCTGGCGCCCTTCCTAGATGCTCTCCCTA 638
Db      436 CTGGGAATGAGAGAGGTCCCACTTGCCTCTCTGTAAGCAGTGGCCATCCTTC 495
Qy      639 GGGAAATG--GTCCAGACAGAGTGGCAGTTCAATGTGGGGGCTGATGTTGTGCTG 696
Db      496 CCTAGGCGAGAAATCCCAAGTGTCTACTTTCCTGGAGGGGCTCATGTTTATC---- 551
Qy      697 GAGGAGACGGCTTACATGTTGTTCTGTAGAAATAAATACTGAGCTA 745
Db      552 -----TGGTCTTAATGTTTGTCTTACTACAGAAATAAATACTGAGTA 594

```

Search completed: February 13, 2004, 14:23:07
Job time : 358 secs

QY	510	CTCCAGAGAGCTTCCCGTTGATGTGGATACAGCAGTCTTCTCG-CCCGTTTTTTGGC	568
Db	674	CTTACACAGCTTTCCTCGTGTGTGGGACACATGACGCGTTTGGGCCCCCGTTTCCGC	733
QY	569	TCAGCCTCCTCAAGGGCAGAGGCGCTTAAGCCAGCCTGAGCCCTTCTTAGTCATGCC	628
Db	734	TAAGCCCTGCGCTGGGCAATGAGCAGAGTCCCATTTGCGTCTCTTAGCAGTGGC	793
QY	629	TCCTCCCTTAGGGATG--GTCCAGCACAAGATGGCCAGTTCATTGTGGGGCCTGATTG	686
Db	794	TCCATCTTCCCTTAGGCGAGAAATCCCAACAATCTACTTTCCTCGGAGGGGCTCAATGT	853
QY	687	TTTGTGCTGTGAGAGAGACGGCTTACATCTTTGTTTCTGTAGAAATTAACGTAGCTAC	746
Db	854	TTTATC-----TGGTCTTAATTTTGTACTACAGAAATTAACGTGCGCTAC	903
QY	747	GA AAAA 752	
Db	904	TA AAAA 909	

RESULT 13
US-10-105-891-47
; Sequence 47, Application US/10105891
; Publication No. US20030073099A1

```

GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenchua
APPLICANT: Zhou, Ping
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Zhao, Qing A.
APPLICANT: Xue, Aidong J.
APPLICANT: Zhang, Jie
APPLICANT: Wehrman, Tom
APPLICANT: Wang, Jian-Rui
APPLICANT: Drmanac, Radjoe T.
TITLE OF INVENTION: No. US20030073099A1el Nucleic Acids and
FILE REFERENCE: 79ICIP2A
CURRENT APPLICATION NUMBER: US/10/105,891
CURRENT FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER: 09/666,317
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 09/552,929
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 91
SOFTWARE: pf_fl_genes Version 2.0
SEQ ID NO 47
LENGTH: 916
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (315)..(636)
US-10-105-891-47

```

	Query Match	17.0%	Score 127.6;	DB 15;	Length 916;
	Best Local Similarity	55.8%;	Pred.No. 7.3e-26;		
	Matches	405;	Conservative	0;	Mismatches 284; Indels 37; Gaps 7.
QY	30	GAGCGGGCAGAGGCTCCGAGCCATGACAGGCCGAAGCCGGGGGCACAGGGGGTTCGACG	89		
Db	218	GGGGGGGGGCTTCGGGAAAGAGCCACAAGAGGGGGGGCGTAGAACCCTGCCTTCGGGG	277		
QY	90	GGCGATTGTATGCGCCCAAGAGGCCCTTGCAATTCCTGATGCGCCAGGGGGCAATCTGGC	149		
Db	278	GTGAGATGTGAGCGCGCGCGCG--CGATCATGCGGAGCGCGGATGACAGACGACAGCGGA	335		
QY	150	GGCCCCAGAGAGCGGGTGGCCACGGGGCGGCAGAAGTCCCCTGGGGGGCGCAGGGGCACGACAG	209		
Db	336	GGCGCTGACGGGGGATATGGCCGGGGTATGCCACACTGCTCCGGGGGGGGGTGACACAGCG	395		

210 GCTCCGAGGACCGGAGAGAGACGCCCGCGCGAGGATCCGATGACGCGCGCTTACAGGGCTG 269
 396 GCAGCTCCGGCCGGATGAGACTCCCGCAGCGACGCGCAGG-----TCCGGCAG 445
 270 AATGATGCTGCAGATGCGGGGCCCGAGGGGCCGAGAGACCGCGCTTGTAAGTTTACCTC 329
 446 A-----GAGCGCGCGCTGCGCCAGGGGGTACAGAAATGCGGCGCGCAGATATTCACCTC 459
 330 GCCATGCTTTGCGCAGCACCCATGGAAGCAGAGCTGCGCCGACAGACCTGSCCCAGAT 389
 500 AGGTGCTTTCCCGACCCCCCTTGGAGGGGGAATTCGCCCATGGGTCCCTGCACCGAGT 559
 390 GCCCCACCGGTTCCGTGCGAGGGGTCTTTGAAGAGTTCACTGTGTCCGCAACA 449
 560 GCCGAGCCCA-----CCMAAGGATGTTGGAAAGATTCACAGTAGAGTGCAGAGATC 613
 450 CTGACTATCCGACTGACTGCTGCAGACACCGCCCACTGACGCTCTCATCAGTCCGT 509
 614 CTGTGTGTCGCTGAAAGCTGAAGACTGCTGCGCTGCTCCGAAATTCGATCACTTT 673
 510 CTCACAGACTTTCCCTGTGATGTGATTCACGCAATGCTTTCTG-CCGCTGTTTGGC 568
 674 CTTGACACACTTTCCCTGTGTGTGGGACATGACGAGCTTTGGGCCCCCGCTTCCGC 723
 569 TCAGCTCCCTCAGGGCAGAGGCGTAAAGCCAGCTGCGGCCCTCTCCTAGGTCAAGCC 628
 734 TAAGCTGGGCTGGGGAAATGAGAGAGTCCACTTTGGCTCTCTTGAAGCAATGCG 793
 629 TCTCTCCCTAAGGAATG--GTCCAGCAAGATGCCAGTTCAATTGTGGGGGCTGAATTG 666
 794 TCCATCTCTTCCCTAAGGACAGAAATCCACAGTTGCACTTCCCTGGAGAGGCTCATGT 853
 687 TTGTGCTGAGAGAGACGCTTACATGTTGTTTCTGTGAAAAATAAACTGAGCTAC 746
 854 TTTATC-----TGGTCTTAAATGTTTGTGTTACTACAGAAAAATAAACTGCGCTAC 903
 747 GAAAAA 752
 904 TAAAAA 909

RESULT 14
US-09-880-107-3826
; Sequence 3826, Application US/09880107
; Patent No. US20020142981A1

```

1  GENERAL INFORMATION:
2  APPLICANT: Horne, Darci T.
3  APPLICANT: Vockley, Joseph G.
4  APPLICANT: Scherf, Iwe
5  APPLICANT: Gene Logic, Inc.
6  TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
7  FILE REFERENCE: 44921-5028-NO
8  CURRENT APPLICATION NUMBER: US/09/880,107
9  CURRENT FILING DATE: 2001-06-14
10 PRIOR APPLICATION NUMBER: US 60/211,379
11 PRIOR FILING DATE: 2000-06-14
12 PRIOR APPLICATION NUMBER: US 60/237,054
13 PRIOR FILING DATE: 2000-10-02
14 NUMBER OF SEQ ID NOS: 3950
15 SOFTWARE: PatentIn Ver. 2.1
16 SEQ ID NO 3826
17 LENGTH: 599
18 TYPE: DNA
19 ORGANISM: Homo sapiens
20 FEATURE:
21 OTHER INFORMATION: Genbank Accession No. US20020142981A1 X928856
22 US-09-880-107-3826

```

Query Match	15.6%	Score 117.4;	DB 10;	Length 599;
Best Local Similarity	59.5%;	Pred. No. 4.9e-23;		
Matches 279;	Conservative 0;	Mismatches 171;	Indels 19;	Gaps 4
OY	280	GCAAGATCGGGGCCAGGGGGCCGAGAGACCGCCCTGTAGATTCCTACCTGCCATGCGCTT	339	

LENGTH: 993
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-157-031-87

Query Match 55.0%; Score 413.4; DB 15; Length 993;
 Best Local Similarity 71.5%; Pred. No. 5.6e-106;
 Matches 677; Conservative 0; Mismatches 41; Indels 229; Gaps 1;

35 GGGCAGAGGCTCCGAGCCATGACAGCCGAAAGCCGGGCAAGAGGGTTCAGAGGCGA 94
 9 GGGCAGAGGCTCCGAGCCATGACAGCCGAAAGCCGGGCAAGAGGGTTCAGAGGCGA 68
 95 TGCTATGAGCCAGAGAGGCTTCGATTCGATGAGCCCAAGGGGCAATGCTGCGGCCC 154
 69 TGCTATGAGCCAGAGAGGCTTCGATTCGATGAGCCCAAGGGGCAATGCTGCGGCCC 128
 155 AGGAGAGGCGGGTGCACAGGGGCGAGAGGTCCCGGGGCGAGGGGCGAGAGAGGCGCTC 214
 129 AGGAGAGGCGGGTGCACAGGGGCGAGAGGTCCCGGGGCGAGGGGCGAGAGAGGCGCTC 188
 215 GGGGCGGGAGAGAGGCGCGCGGCGGTCGCGATGCGCGCGGCTTCAGGGCTGAGATG 274
 189 GGGGCGGGAGAGAGGCGCGCGGCGGTCGCGATGCGCGCGGCTTCGCGCGAGATG 248
 275 ATGCTGAGATGCGGGGCGAGGGGCGCGAGCGCGCTGCTGAGTTCATCTCCGCAT 334
 249 AAGGTGCGCGGCGAGAGGCGCGAGCGCGCTGCTGAGTTCATCTCCGCAT 308
 335 GCCTTCGCGACACCCATGGAAGCAGAGTGGCCCGCAGAGCCTGCGCCAGATGCCCC 394
 309 GCCTTCGCGACACCCATGGAAGCAGAGTGGCCCGCAGAGCCTGCGCCAGATGCCCC 368
 395 ACCGCTCCCGTCCGAGGGGTCTCTGAGAGATTCATCTGTCTCCGCAACATCTGAC 454
 369 ACCGCTCCCGTCCGAGGGGTCTCTGAGAGATTCATCTGTCTCCGCAACATCTGAC 428
 455 TAT----- 457
 429 TATGTCATGTCGGAGACAGAGACAGGGAGCGCTGGGGAGATGAGGGTGTGGTGGGG 488
 458 ----- 457
 489 GCTGGATCCGCTCCCGAGAGGGCAAGAACTAGAGATTCAGAGACCCCAACAA 548
 458 ----- 457
 549 GGTCTCAGAAACAGAGACTGTACACAGGCCCGCCGCCACCCAGAGGAGCCAGAGAGA 608
 458 -----CGACTGA 465
 609 TGGGTGACAGAGGTGTCGCTTAAATGATGTTCTGCCCCCTCAATTTAGCCGACGA 668
 466 CTGCTGACAGACCAAGCCCAACTGACGCTCTCATAGCTCCTGTCTCAGAGAGCTTCC 525
 663 CTGCTGACAGACCAAGCCCAACTGACGCTCTCATAGCTCCTGTCTCAGAGAGCTTCC 728
 526 TGTATGATGATCAACGAGTCTTTTGCCGCTTTTGGCTCAGGCTCCCTCAGGGC 585
 729 TGTATGATGATCAACGAGTCTTTTGCCGCTTTTGGCTCAGGCTCCCTCAGGGC 788
 586 AAGAGCGCTAAGCCCAAGCTGAGCGCCCTTCTAGATGATGCTCTCCCTAAGAGATG 645
 789 AAGAGCGCTAAGCCCAAGCTGAGCGCCCTTCTAGATGATGCTCTCCCTAAGAGATG 848
 646 GTCCAGACAGAGTGCAGATTCATTTGAGGGGCTGATTTGTTGCTGAGAGAGAC 705
 849 GTCCAGACAGAGTGCAGATTCATTTGAGGGGCTGATTTGTTGCTGAGAGAGAC 908
 706 GGTCTACATGTTGTTCTGTAGAAATTAACCTAGAGCAAGAAA 752
 909 GGTCTACATGTTGTTCTGTAGAAATTAACCTAGAGCAAGAAA 955

RESULT 12

US-10-125-237-47
 Sequence 47, Application US/10125237
 Publication No. US20030022329A1
 GENERAL INFORMATION:
 APPLICANT: Tang, Y. Tom
 APPLICANT: Liu, Chenghua
 APPLICANT: Zhou, Ping
 APPLICANT: Aoudi, Vinod
 APPLICANT: Ren, Feiyan
 APPLICANT: Zhao, Qing A.
 APPLICANT: Xue, Aidong J.
 APPLICANT: Zhang, Jie
 APPLICANT: Wehrman, Tom
 APPLICANT: Wang, Jian-Rui
 APPLICANT: Drmanac, Radote T.
 TITLE OF INVENTION: No. US20030022329A1el Nucleic Acids and
 TITLE OF INVENTION: Polypeptides
 FILE REFERENCE: 79ICP2ADIV
 CURRENT APPLICATION NUMBER: US/10/125,237
 PRIOR FILING DATE: 2002-04-17
 PRIOR APPLICATION NUMBER: 09/668,317
 PRIOR FILING DATE: 2000-09-22
 PRIOR APPLICATION NUMBER: 09/552,929
 NUMBER OF SEQ ID NOS: 91
 SOFTWARE: pc_fl_genes Version 2.0
 SEQ ID NO 47
 LENGTH: 916
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (319)..(636)
 US-10-125-237-47

Query Match 17.0%; Score 127.6; DB 15; Length 916;
 Best Local Similarity 55.8%; Pred. No. 7.3e-26;
 Matches 405; Conservative 0; Mismatches 284; Indels 37; Gaps 7;

30 GAGCCGGGCAAGAGCTCCGAGCCATGACAGCCGAAAGCCGGGCAAGAGGGGTTGACG 89
 218 GGGGCGGGGCTTGAGGAAGAGCCAGAGAGGCGGGCGTAGACCTGCGCTTCGGGG 277
 90 GGGATGCTGATGAGCCCAAGAGGCTTCGATTCGATGAGCCCAAGGGGCAATGCTGAC 149
 278 GTGAGTGGAGAGGGGCGGGCGG--CGTCAATGCGGAGCGGATGCAACGAGCGGA 335
 150 GGGCCAGAGAGGGGCGGTGCACAGGGCGCAGAGTCCCGGGGCGCAGAGGAGCAAG 209
 336 GGGCGTGAACGGCGGGATGAGCCGGGGTGCACAGCTGCGGGGCGTGAACAGGCC 395
 210 GCTCGGGGCGGGGAGAGAGCGCCCGGGGTCGATGAGCGGCGGCTTCAGGGCTG 269
 396 GCACTTCGGCGGTGAGTCTCCCAAGCGCACGGCCAGG-----TCCGGGAG 445
 270 AATGATGCTGAGATGCGGGGCGAGGGGCGGAGAGCGGCTGTTAGTTTCACTC 329
 446 A-----GACGCGCGCTTGGCGCAAGGGGTCAAGAAACCGCGCACATATTCAACCTC 499
 330 GCCATGCTTTGGGACACCCATGGAAGAGAGTGGCCCGAGAGGCTTCGCGCCAGAT 389
 500 AGGTGCTTTCCGACCCCTTGAAGGCGAAATGCCCATAGAGTCCCTGAGCAACAGAT 559
 390 GCCCACCGCTTCGTCAGAGGAGTCTTGAAGAGTCACTGTGTCCGCAACATA 449
 560 GCCAGGCCCA-----CCAAAGGGGTGTGGAAGAGATCAAGTGAAGTGGAGATC 613
 450 CTGACTATCCGACTGACTGTGAGACCAACGGCCAACTGACACTCTCCATCAGCTCTGT 509
 614 CTGCTGCTCGCTGGAAGAGTGTGCTGCTGCTGCTGCAATTCCTGTCATCAACTTT 673

Db 241 AGGGGGCCGGAGAGCCGCTGTTAGTTTACTCTCCGACCTCTTCCGACACCCATG 300
 QY 354 GAGCAGAGCTGGCCCGCAGAGAGCTGGCCCGCAGATGCCCCACCGCTTCCCGTCCAGGG 413
 Db 301 GAGCAGAGCTGGCCCGCAGAGAGCTGGCCCGCAGATGCCCCACCGCTTCCCGTCCAGGG 360
 QY 414 GTGCTCTGAGAGAGTTGACTGTGTGCGGCAACATAGTACTATCCGACTGCTGCTGCA 473
 Db 351 GTGCTCTGAGAGAGTTGACTGTGTGCGGCAACATAGTACTATCCGACTGCTGCTGCA 420
 QY 474 GACCAAGCCCACTGACCTCTTCATCATGCTCTGTCTCCAGAGCTTTCCTGTGATG 533
 Db 421 GACCAAGCCCACTGACCTCTTCATCATGCTCTGTCTCCAGAGCTTTCCTGTGATG 480
 QY 534 TGGATCAGGAGTGTCTTCTGCGCGGTGTTTGGCTGACGCTCCCTCAGAGGCAAGGCGC 593
 Db 481 TGGATCAGGAGTGTCTTCTGCGCGGTGTTTGGCTGACGCTCCCTCAGAGGCAAGGCGC 540
 QY 594 TAAGC 598
 Db 541 GAGCG 545

RESULT 10

US-10-295-027-387
 ; Sequence 387, Application US/10295027
 ; Publication No. US20030232350A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Afar, Daniel
 ; APPLICANT: Aziz, Natasha
 ; APPLICANT: Gineberg, Wendy M.
 ; APPLICANT: Glen, Kurt C.
 ; APPLICANT: Glynn, Richard
 ; APPLICANT: Hevezl, Peter A.
 ; APPLICANT: Mack, David H.
 ; APPLICANT: Murray, Richard
 ; APPLICANT: Watson, Susan R.
 ; APPLICANT: Eos Biotechnology, Inc.
 ; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
 ; FILE REFERENCE: 018501-012500US
 ; CURRENT APPLICATION NUMBER: US/10/295,027
 ; CURRENT FILING DATE: 2002-11-13
 ; PRIOR APPLICATION NUMBER: US 09/663,733
 ; PRIOR FILING DATE: 2000-09-15
 ; PRIOR APPLICATION NUMBER: US 60/350,666
 ; PRIOR FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 60/335,394
 ; PRIOR FILING DATE: 2001-11-15
 ; PRIOR APPLICATION NUMBER: US 60/332,464
 ; PRIOR FILING DATE: 2001-11-21
 ; PRIOR APPLICATION NUMBER: US 60/334,393
 ; PRIOR FILING DATE: 2001-11-29
 ; PRIOR APPLICATION NUMBER: US 60/340,376
 ; PRIOR FILING DATE: 2001-12-14
 ; PRIOR APPLICATION NUMBER: US 60/347,211
 ; PRIOR FILING DATE: 2002-01-08
 ; PRIOR APPLICATION NUMBER: US 60/347,349
 ; PRIOR FILING DATE: 2002-01-10
 ; PRIOR APPLICATION NUMBER: US 60/355,250
 ; PRIOR FILING DATE: 2002-02-08
 ; PRIOR APPLICATION NUMBER: US 60/356,714
 ; PRIOR FILING DATE: 2002-02-13
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 1386
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 387
 ; LENGTH: 608
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-295-027-387

Query Match

57.7%; Score 434.2; DB 12; Length 608;

Best Local Similarity 79.4%; Pred. No. 7,6e-112;
 Matches 590; Conservative 0; Mismatches 18; Indels 135; Gaps 1;

QY 3 CCTGTGGGCGCCGACCTTCTCTGTGAGAGCGGCGAGGCTCCGAGCCATGAGGCG 62
 Db 1 CCTGTGGGCGCCGACCTTCTCTGTGAGAGCGGCGAGGCTCCGAGCCATGAGGCG 60
 QY 63 GAAAGCCGGGCGACAGGGGGTTTGACGGGCGATGCTGATGCTCCAGAGGCGCTTGATT 122
 Db 61 GAAAGCCGGGCGACAGGGGGTTTGACGGGCGATGCTGATGCTCCAGAGGCGCTTGATT 120
 QY 123 CTTGATGGCCCAAGGGGGCAATGCTGCGGCGCCAGAGAGAGCGGATGCCAGCGGCGAG 182
 Db 121 CTTGATGGCCCAAGGGGGCAATGCTGCGGCGCCAGAGAGAGCGGATGCCAGCGGCGAG 180
 QY 183 GGTCCCGGGCGCGAGGGCGAGAGGCTTGGGGCGGGCGAGAGGCGCCCGCGGGGT 242
 Db 181 GGTCCCGGGCGCGAGGGCGAGAGGCTTGGGGCGGGCGAGAGGCGCCCGCGGGGT 240
 QY 243 CCGCATGGGCGGCGGCTTCAAGGGCTGAATGATGCTGCAATGGGGGCGAGGGGCGCG 302
 Db 241 CCGCATGGGCGGCTTCAAGGGCTTCTGCGAGATGATGATGCTGCGGCGCGAGGGGCGCG 300
 QY 303 GAGAGCGGCGCTTGTGATTTACTCTGCGCATGCTTTCGCGACACCATGAGACAGAG 362
 Db 301 GAGAGCGGCGCTTGTGATTTACTCTGCGCATGCTTTCGCGACACCATGAGACAGAG 319
 QY 363 CTGGCCCGGAGAGGCTGGCCCGACAGATGCCCCACCGCTTCCGTGCGAGGGGTGCTTCTG 422
 Db 320 -----TCCGATGACGCTGCGAGACCAACCGC 319
 QY 423 AAGGAGTTCACTGTGTGCGGCAACATAGTACTATCCAGTACTGCTGAGACACCGC 482
 Db 320 -----TCCGATGACGCTGCGAGACCAACCGC 345
 QY 483 CAAGTCAAGCTTCCATCAAGCTCTGTCTTCCAGAGCTTTCCTGTGATGATGATCAG 542
 Db 346 CAAGTCAAGCTTCCATCAAGCTCTGTCTTCCAGAGCTTTCCTGTGATGATGATCAG 405
 QY 543 CAGTGTCTTCTGCGGCTGTTTGTGCTGAGCTCCCTCAGAGGCGAGGCGCTTAAAGCCAG 602
 Db 406 CAGTGTCTTCTGCGGCTGTTTGTGCTGAGCTCCCTCAGAGGCGAGGCGCTTAAAGCCAG 465
 QY 603 CTTGGGCGCCCTTCTTCAAGTCAAGCTCTTCCCTCAGAGGATGCTCCAGACAGATGGC 662
 Db 466 CTTGGGCGCCCTTCTTCAAGTCAAGCTCTTCCCTCAGAGGATGCTCCAGACAGATGGC 525
 QY 663 CAGTCAATGTGGGGCGCGATGTTGTGCTGAGAGAGAGCGCTTACATGTTTGT 722
 Db 526 CAGTCAATGTGGGGCGCGATGTTGTGCTGAGAGAGAGCGCTTACATGTTTGT 585
 QY 723 CTGTAGAAATTAATTAAGCTGAGCTA 745
 Db 586 CTGTAGAAATTAATTAAGCTGAGCTA 608
 RESULT 11
 US-10-157-031-87
 ; Sequence 87, Application US/10157031
 ; Publication No. US20030108890A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baranova, A. V.
 ; APPLICANT: Yankovsky, N. K.
 ; APPLICANT: Kozlov, A. P.
 ; APPLICANT: Lobashov, A. V.
 ; APPLICANT: Kutkovskaya, L. L.
 ; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
 ; FILE REFERENCE: 2760-103
 ; CURRENT APPLICATION NUMBER: US/10/157,031
 ; CURRENT FILING DATE: 2002-05-30
 ; NUMBER OF SEQ ID NOS: 415
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 87

```

QY 543 CAGTCTTTCGCCCCGTTTGTGCTAGACCTCCCTGAGGCGCAGAGGCGCTAAAGCCGAC 602
DB 578 CAGTCTTTCGCCCCGTTTGTGCTAGACCTCCCTGAGGCGCAGAGGCGCTAAAGCCGAC 637
QY 603 CCGGCGCCCTTCTAGTGCATGCTCTCCCTTAAGGAAATGATCCAGACAGAGTGGC 662
DB 638 CCGGCGCCCTTCTAGTGCATGCTCTCCCTTAAGGAAATGATCCAGACAGAGTGGC 697
QY 663 CAGTCTTTCGCCCCGTTTGTGCTAGACCTCCCTGAGGCGCAGAGGCGCTAAAGCCGAC 722
DB 698 CAGTCTTTCGCCCCGTTTGTGCTAGACCTCCCTGAGGCGCAGAGGCGCTAAAGCCGAC 757
QY 723 CTGTAGAAAATAAACTAGCTA 745
DB 758 CTGTAGAAAATAAACTAGCTA 780

```

```

RESULT 8
US-10-146-473-29
; Sequence 29, Application US/10146473
; Publication No. US2003010888A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: 100461/70130 (IRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29
; LENGTH: 750
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-146-473-29

```

```

Query Match 90.4%; Score 679.6; DB 15; Length 750;
Best Local Similarity 94.1%; Pred. No. 1.2e-180;
Matches 706; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```

```

QY 1 ATCTGTGAGGCTCTGACCTTCTCTGAGAGCCGAGGCTCCGAGCCATGACAG 60
DB 1 ATCTGTGAGGCTCTGACCTTCTCTGAGAGCCGAGGCTCCGAGCCATGACAG 60
QY 61 CCGAAGGCGCGGCGCAAGGCGGTTGAGCGGCGATGATGAGGCGCCAGAGGCGCTTGCA 120
DB 61 CCGAAGGCGCGGCGCAAGGCGGTTGAGCGGCGATGATGAGGCGCCAGAGGCGCTTGCA 120
QY 121 TTCTGTATGAGCCCAAGGCGGCAATGCTGCGCGCCAGAGAGGCGGATGCCA 180
DB 121 TTCTGTATGAGCCCAAGGCGGCAATGCTGCGCGCCAGAGAGGCGGATGCCA 180
QY 121 TTCTGTATGAGCCCAAGGCGGCAATGCTGCGCGCCAGAGAGGCGGATGCCA 180
DB 121 TTCTGTATGAGCCCAAGGCGGCAATGCTGCGCGCCAGAGAGGCGGATGCCA 180
QY 181 GAGGTCCCGCGGCGCGAGGCGGCAAGGCGGCTCGGCGCGGAGAGAGGCGCGCGG 240
DB 181 GAGGTCCCGCGGCGCGAGGCGGCAAGGCGGCTCGGCGCGGAGAGAGGCGCGCGG 240
QY 241 GTCCGATGAGCGCGCGCTTCAAGGCTGATGATGATGCTGAGATGCGGAGGCGC 300
DB 241 GTCCGATGAGCGCGCGCTTCAAGGCTGATGATGATGATGCTGAGATGCGGAGGCGC 300
QY 301 CGGAGAGCGCGCTGAGTGTCTAAGTCTGAGGCGGCTTTCGCGCAAGGAGAGAG 360
DB 301 CGGAGAGCGCGCTGAGTGTCTAAGTCTGAGGCGGCTTTCGCGCAAGGAGAGAG 360
QY 361 AGCTGAGCGCGAGAGCTGAGGCGGCGGAGAGGCGGCTTTCGCGCAAGGAGAG 420
DB 361 AGCTGAGCGCGAGAGCTGAGGCGGCGGAGAGGCGGCTTTCGCGCAAGGAGAG 420

```

```

QY 421 TGAAGAGTTCACTGTGTCCGCGCAACATCTAGCTATCCGACTGATGTGAGACCAAC 480
DB 421 TGAAGAGTTCACTGTGTCCGCGCAACATCTAGCTATCCGACTGATGTGAGACCAAC 480
QY 481 GCCAATGAGAGCTCTCATCAGTCTGTCTGAGAGAGCTTCCCTGTATGATGAGATCA 540
DB 481 GCCAATGAGAGCTCTCATCAGTCTGTCTGAGAGAGCTTCCCTGTATGATGAGATCA 540
QY 541 CGCAGTCTTTCGCGCGGTTTGTGCTAGAGCTCTCCCTAAGGAGAGAGGCGCTAAGCC 600
DB 541 CGCAGTCTTTCGCGCGGTTTGTGCTAGAGCTCTCCCTAAGGAGAGAGGCGCTAAGCC 600
QY 601 AGCCTGAGCGCCCTTCTAGTGCATGCTCTCCCTAAGGAGAGAGGCGCTAAGCC 660
DB 601 AGCCTGAGCGCCCTTCTAGTGCATGCTCTCCCTAAGGAGAGAGGCGCTAAGCC 660
QY 661 GCCAGTTCAATGTGAGGAGGCTGATGTTGTGCTGAGAGAGAGAGGCGCTTACATGTTGT 720
DB 661 GCCAGTTCAATGTGAGGAGGCTGATGTTGTGCTGAGAGAGAGAGGCGCTTACATGTTGT 720
QY 721 TTCTGTAGAAAATAAACTAGCTAGCAAA 750
DB 721 TTCTGTAGAAAATAAACTAGCTAGCTAGCAAA 750

```

```

RESULT 9
US-09-821-883-28
; Sequence 28, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vlodavich, Damiir
; APPLICANT: Gradiš, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: NY-ESO-1C tumor antigen
US-09-821-883-28

```

```

Query Match 71.8%; Score 540.2; DB 9; Length 1191;
Best Local Similarity 99.4%; Pred. No. 1.6e-141;
Matches 542; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 54 ATGCAAGGCGGAGCGCGGCGCAAGGCGGTTGAGCGGCGGATGCTGATGAGGCGCCAGAGAGC 113
DB 1 ATGCAAGGCGGAGCGCGGCGCAAGGCGGTTGAGCGGCGGATGCTGATGAGGCGCCAGAGAGC 113
QY 114 CTTGCGCAATTCCTATGAGGCGGCGCAATGCTGCGGCGCCAGAGAGGCGGATGCAAG 173
DB 114 CTTGCGCAATTCCTATGAGGCGGCGCAATGCTGCGGCGCCAGAGAGGCGGATGCAAG 173
QY 174 GCGCGAGAGGTCCTCGGCGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 233
DB 174 GCGCGAGAGGTCCTCGGCGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 233
QY 234 CCGCGGAGGTCCTCGGCGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 293
DB 234 CCGCGGAGGTCCTCGGCGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 293
QY 294 AGGAGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 353
DB 294 AGGAGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 353

```

Query Match 99.6%; Score 749; DB 15; Length 755;
 Best Local Similarity 100.0%; Pred. No. 4,4e-200; Indels 0; Gaps 0;
 Matches 749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 4 CTCGTGGGCCCCGACCTTCTCTGAGAGCCGCGGAGAGGCTCCGAGCATTGAGGCGC 63
DB 1 CTCGTGGGCCCCGACCTTCTCTGAGAGCCGCGGAGAGGCTCCGAGCATTGAGGCGC 60
QY 64 AAGGCGGCGGCGACAGGCGGCTTGAATGGGCTGATGCTGATGCTGATGCTGATGCT 123
DB 61 AAGGCGGCGGCGACAGGCGGCTTGAATGGGCTGATGCTGATGCTGATGCTGATGCT 120
QY 124 CTGATGCGCCAGAGGCGGCGATGCTGAGGCGCCAGAGAGGCGGCGGCGGCGGCGG 183
DB 121 CTGATGCGCCAGAGGCGGCGATGCTGAGGCGCCAGAGAGGCGGCGGCGGCGGCGG 180
QY 184 GTCCCGGCGGCGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 243
DB 181 GTCCCGGCGGCGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
QY 244 CGCATGCGGCGGCGGCGGCTTCAAGGCGGCTGATGATGCTGATGCTGATGCTGATGCT 303
DB 241 CGCATGCGGCGGCGGCGGCTTCAAGGCGGCTGATGATGCTGATGCTGATGCTGATGCT 300
QY 304 AAGCGCGGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 363
DB 301 AAGCGCGGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 360
QY 364 TGGCGCGGCGGAGAGCGTGGCCAGATGCGCCAGATGCGCCAGATGCGCCAGATGCGCC 423
DB 361 TGGCGCGGCGGAGAGCGTGGCCAGATGCGCCAGATGCGCCAGATGCGCCAGATGCGCC 420
QY 424 AGGATGCTACTGTGTCCGCAACAATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 483
DB 421 AGGATGCTACTGTGTCCGCAACAATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 480
QY 484 AACTGCGAGCTCTCCATGAGCTCTGTCTGCGAGAGCTTCCCTGTTGATGATGATGAT 543
DB 481 AACTGCGAGCTCTCCATGAGCTCTGTCTGCGAGAGCTTCCCTGTTGATGATGATGAT 540
QY 544 AGTGTCTTCTGCGCGTGTGTTTGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 603
DB 541 AGTGTCTTCTGCGCGTGTGTTTGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 600
QY 604 CTGCGCGGCGGCTTCTGAGCTGATGCTGCTGCTGCTGAGGAGATGCTGAGAGAGCT 663
DB 601 CTGCGCGGCGGCTTCTGAGCTGATGCTGCTGCTGCTGAGGAGATGCTGAGAGAGCT 660
QY 664 AGTTCATGTTGCGGCGGCTGATGTTGTTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 723
DB 661 AGTTCATGTTGCGGCGGCTGATGTTGTTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 724 TGTAGAAAATAAACTGAGCTAGCAAAAA 752
DB 721 TGTAGAAAATAAACTGAGCTAGCAAAAA 749

```

RESULT 7
 US-10-295-027-385
 Sequence 385; Application US/10295027
 Publication No. US2003023350A1

GENERAL INFORMATION:

APPLICANT: Afar, Daniel
 APPLICANT: Aziz, Natasha
 APPLICANT: Ginsberg, Wendy M.
 APPLICANT: Glush, Kurt C.
 APPLICANT: Glynn, Richard
 APPLICANT: Hevezi, Peter A.
 APPLICANT: Mack, David H.
 APPLICANT: Murray, Richard
 APPLICANT: Watson, Susan R.
 APPLICANT: Eos Biotechnology, Inc.
 TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and

TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
 FILE REFERENCE: 018501-012500US
 CURRENT APPLICATION NUMBER: US/10/295,027
 CURRENT FILING DATE: 2002-11-13
 PRIOR APPLICATION NUMBER: US 09/663,733
 PRIOR FILING DATE: 2000-09-15
 PRIOR APPLICATION NUMBER: US 60/350,666
 PRIOR FILING DATE: 2001-11-13
 PRIOR APPLICATION NUMBER: US 60/335,394
 PRIOR FILING DATE: 2001-11-15
 PRIOR APPLICATION NUMBER: US 60/332,464
 PRIOR FILING DATE: 2001-11-21
 PRIOR APPLICATION NUMBER: US 60/334,393
 PRIOR FILING DATE: 2001-11-29
 PRIOR APPLICATION NUMBER: US 60/340,376
 PRIOR FILING DATE: 2001-12-14
 PRIOR APPLICATION NUMBER: US 60/347,211
 PRIOR FILING DATE: 2002-01-08
 PRIOR APPLICATION NUMBER: US 60/347,349
 PRIOR FILING DATE: 2002-01-10
 PRIOR APPLICATION NUMBER: US 60/355,250
 PRIOR FILING DATE: 2002-02-08
 PRIOR APPLICATION NUMBER: US 60/356,714
 PRIOR FILING DATE: 2002-02-13
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 1386
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 385
 LENGTH: 780
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-295-027-385

Query Match 98.8%; Score 743; DB 12; Length 780;
 Best Local Similarity 100.0%; Pred. No. 2.1e-198; Indels 0; Gaps 0;
 Matches 743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 3 CCTGTGGGCCCCGACCTTCTCTGAGAGCCGCGGAGAGGCTCCGAGCATTGAGGCGC 62
DB 38 CCTGTGGGCCCCGACCTTCTCTGAGAGCCGCGGAGAGGCTCCGAGCATTGAGGCGC 97
QY 63 GAAAGCGGCGGCGGAGAGGCGGCTTCAAGGCGGCTGATGCTGATGCTGATGCTGATGCT 122
DB 98 GAAAGCGGCGGCGGAGAGGCGGCTTCAAGGCGGCTGATGCTGATGCTGATGCTGATGCT 157
QY 123 CCTGATGCGCCAGAGGCGGCAATGCTGCGGCGCCAGAGAGAGCGGCTGAGAGAGAG 182
DB 158 CCTGATGCGCCAGAGGCGGCAATGCTGCGGCGCCAGAGAGAGCGGCTGAGAGAGAG 217
QY 183 GGTTCGCGGCGGCGGAGAGGCGGCTTGGGCGGCGGAGAGAGAGAGAGAGAGAGAGAG 242
DB 218 GGTTCGCGGCGGCGGAGAGGCGGCTTGGGCGGCGGAGAGAGAGAGAGAGAGAGAGAG 277
QY 243 CCGCATGCGGCGGCGGCTTCAAGGCGGCTGATGCTGATGCTGATGCTGATGCTGATGCT 302
DB 278 CCGCATGCGGCGGCGGCTTCAAGGCGGCTGATGCTGATGCTGATGCTGATGCTGATGCT 337
QY 303 GAGAGCGGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 362
DB 338 GAGAGCGGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 397
QY 363 CTGCGCGGCGGAGAGCGTGGCCAGAGATGCGCCAGAGCTTCCGTCGAGAGGCGGCTTCT 422
DB 398 CTGCGCGGCGGAGAGCGTGGCCAGAGATGCGCCAGAGCTTCCGTCGAGAGGCGGCTTCT 457
QY 423 AAGGATGCTACTGTGTCCGCAACAATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 482
DB 458 AAGGATGCTACTGTGTCCGCAACAATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 517
QY 483 CAAGCGAGCTCTCATAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 542
DB 518 CAAGCGAGCTCTCATAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 577

```


Db 481 GCCAATGCAAGCTCTCCATCAGCTCCTGTCTCCAGAGCTTCCCTGTGATGTGATCA 540
QY 541 CGCAGTGTCTTCTGCGCCGTGTGTTTGGCTCAGCTTCCCTCAGGCGCAAGCGCTTAAGCCC 600
Db 541 CGCAGTGTCTTCTGCGCCGTGTGTTTGGCTCAGCTTCCCTCAGGCGCAAGCGCTTAAGCCC 600
QY 601 AGCTGAGCGCCCTTCTAGTATGATGCTCTCTCCCTAGAGAAATGGTCCCAAGCAGAGTG 660
Db 601 AGCTGAGCGCCCTTCTAGTATGATGCTCTCTCCCTAGAGAAATGGTCCCAAGCAGAGTG 660
QY 661 GCCAGTTCAATTGTGGGGGCTGATTGTGTCGCTGAGAGAGACGGCTTAATGTTTGT 720
Db 661 GCCAGTTCAATTGTGGGGGCTGATTGTGTCGCTGAGAGAGACGGCTTAATGTTTGT 720
QY 721 TTCTGTAGAAATAAATACTGAGTACGAAAAA 752
Db 721 TTCTGTAGAAATAAATACTGAGTACGAAAAA 752

RESULT 5

US-10-023-182-1
; Sequence 1, Application US/10023182
; Publication No. US20020164665A1
; GENERAL INFORMATION:
; APPLICANT: Stockert, Elisabeth; Jager, Elke;
; Chen, Yao-tseng; Scanlan, Matthew;
; Knuth, Alexander; Old, Lloyd J.
; TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer
; Associated Proteins, Uses Thereof.
; Truncated Forms of NY-ESO-1, and HLA
; Binding Peptides Derived Therefrom
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski, L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 Kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/023,182
; FILING DATE: 17-Dec-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/751,798
; FILING DATE: December 29, 2000
; APPLICATION NUMBER: 09/062,422
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/725,182
; FILING DATE: October 3, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. US20020164665A1man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5466.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3168
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 752 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-023-182-1
Query Match 100.0%; Score 752; DB 14; Length 752;
Best Local Similarity 100.0%; Pred. No. 6,3e-201;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTCGTGGGCTCTGACCTTCTCTGAGAGCGCGGAGAGGCTCCGAGCGCATGACAG 60
Db 1 ATCTCGTGGGCTCTGACCTTCTCTGAGAGCGCGGAGAGGCTCCGAGCGCATGACAG 60
QY 61 CCGAAGGCGCGGACACAGGAGGTTTCAGACGGGCGATGCTGATGAGCCAGAGAGGCTTGGA 120
Db 61 CCGAAGGCGCGGAGACAGGAGGTTTCAGACGGGCGATGCTGATGAGCCAGAGAGGCTTGGA 120
QY 121 TTCTGATGACCCAGAGGAGGAAATGCTGCGGCGCCCAAGAGAGGAGGAGGAGGAGGAG 180
Db 121 TTCTGATGACCCAGAGGAGGAAATGCTGCGGCGCCCAAGAGAGGAGGAGGAGGAGGAG 180
QY 181 GAGGTCCCGGAGGCGCAGAGGAGGCTCGGAGGCGGAGAGAGGAGGAGGAGGAGGAGGAG 240
Db 181 GAGGTCCCGGAGGCGCAGAGGAGGCTCGGAGGCGGAGAGAGGAGGAGGAGGAGGAGGAG 240
QY 241 GTCCGATGAGCGGCGGCGCTTCAAGGCTGAATGATGCTGACATGCGGAGGCGAGGAGG 300
Db 241 GTCCGATGAGCGGCGGCGCTTCAAGGCTGAATGATGCTGACATGCGGAGGCGAGGAGG 300
QY 301 CGAGAGCGCGCTGCTTGAATGCTTACCTGSCATGCTTCCGAGACCAACCATGAAAGCAG 360
Db 301 CGAGAGCGCGCTGCTTGAATGCTTACCTGSCATGCTTCCGAGACCAACCATGAAAGCAG 360
QY 361 AGCTGCGCGCAGAGAGCTTGGCCCAAGAGATGCCACCGCTTCCGAGGAGGAGGAGGAG 420
Db 361 AGCTGCGCGCAGAGAGCTTGGCCCAAGAGATGCCACCGCTTCCGAGGAGGAGGAGGAG 420
QY 421 TGAAGAGTTCACTGATGCTCGGAGCACTAATGATGCTGATGCTGAGCTGAGAGACAC 480
Db 421 TGAAGAGTTCACTGATGCTCGGAGCACTAATGATGCTGATGCTGAGCTGAGAGACAC 480
QY 481 GCCAATGCAAGCTCTCCATCAGCTCCTGTCTCCAGAGCTTCCCTGTGATGTGATCA 540
Db 481 GCCAATGCAAGCTCTCCATCAGCTCCTGTCTCCAGAGCTTCCCTGTGATGTGATCA 540
QY 541 CGCAGTGTCTTCTGCGCCGTGTGTTTGGCTCAGCTTCCCTCAGGCGCAAGCGCTTAAGCCC 600
Db 541 CGCAGTGTCTTCTGCGCCGTGTGTTTGGCTCAGCTTCCCTCAGGCGCAAGCGCTTAAGCCC 600
QY 601 AGCTGAGCGCCCTTCTAGTATGATGCTCTCTCCCTAGAGAAATGGTCCCAAGCAGAGTG 660
Db 601 AGCTGAGCGCCCTTCTAGTATGATGCTCTCTCCCTAGAGAAATGGTCCCAAGCAGAGTG 660
QY 661 GCCAGTTCAATTGTGGGGGCTGATTGTGTCGCTGAGAGAGACGGCTTAATGTTTGT 720
Db 661 GCCAGTTCAATTGTGGGGGCTGATTGTGTCGCTGAGAGAGACGGCTTAATGTTTGT 720
QY 721 TTCTGTAGAAATAAATACTGAGTACGAAAAA 752
Db 721 TTCTGTAGAAATAAATACTGAGTACGAAAAA 752

RESULT 6

US-10-207-655-70
; Sequence 70, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 70
; LENGTH: 755
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-207-655-70

RESULT 3

US-10-364-614-15
; Sequence 15, Application US/10364614
; Publication No. US20030175250A1
; GENERAL INFORMATION:
; APPLICANT: JAGER, Elke
; APPLICANT: KNUTH, Alexander
; APPLICANT: OLD, Lloyd
; APPLICANT: Gujatic, Sacha
; TITLE OF INVENTION: ISOLATED PEPTIDES WHICH BIND TO HLA MOLECULES AND USES THEREOF
; FILE REFERENCE: LRD 5726.1 CIP
; CURRENT APPLICATION NUMBER: US/10/364,614
; PRIOR FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: US 60/355,828
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 15
; LENGTH: 752
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-364-614-15

Query Match 100.0%; Score 752; DB 13; Length 752;
Best Local Similarity 100.0%; Pred. No. 6.3e-201;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTCTGCGGCGCTGACCTTCTCTGAGAGCGGGGAGAGGCTCGGAGCCATGACAG 60
DB 1 ATCTCTGCGGCGCTGACCTTCTCTGAGAGCGGGGAGAGGCTCGGAGCCATGACAG 60
QY 61 CCGAAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAG 120
DB 61 CCGAAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAG 120
QY 121 TTCTGTATGCGCCAGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGG 180
DB 121 TTCTGTATGCGCCAGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGG 180
QY 181 GAGGTCCCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAG 240
DB 181 GAGGTCCCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAG 240
QY 241 GTCCGATGCGCGCGGCTTCAAGGCTGAATGATGCTGATGCGCGCGGCGGCGGCG 300
DB 241 GTCCGATGCGCGCGGCTTCAAGGCTGAATGATGCTGATGCGCGCGGCGGCGGCGG 300
QY 301 CCGAAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAG 360
DB 301 CCGAAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAG 360
QY 361 AGCTGCGCGGAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 420
DB 361 AGCTGCGCGGAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 420
QY 421 TGAAGAGTTCACTGTGTCCGCAACATCTGATTAATGATGCTGATGCTGATGCTGAT 480
DB 421 TGAAGAGTTCACTGTGTCCGCAACATCTGATTAATGATGCTGATGCTGATGCTGAT 480
QY 481 GCCAAGTCAAGTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 540
DB 481 GCCAAGTCAAGTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 540
QY 541 CGCAGTGTCTTTCGCGGCTGTTTGGCTGAGCTTCCCTCAAGGAGAGGCGCTAAGCC 600
DB 541 CGCAGTGTCTTTCGCGGCTGTTTGGCTGAGCTTCCCTCAAGGAGAGGCGCTAAGCC 600
QY 601 AGCTGCGCGGCTTCTGAGTCAATGCTTCCCTCAAGGAGAGGCTTCCCTCAAGGAG 660
DB 601 AGCTGCGCGGCTTCTGAGTCAATGCTTCCCTCAAGGAGAGGCTTCCCTCAAGGAG 660
QY 661 GCCAGTTCATGTTGGGGGCTGATGTTTGTGCTGAGAGAGAGGCTTCCCTGTTGAT 720
DB 661 GCCAGTTCATGTTGGGGGCTGATGTTTGTGCTGAGAGAGAGGCTTCCCTGTTGAT 720

DB 661 GCCAGTTCATGTTGGGGGCTGATGTTTGTGCTGAGAGAGAGGCTTCCCTGTTGAT 720

QY 721 TTCTGTAGAAAATAAACTGAGCTACGAAAA 752
DB 721 TTCTGTAGAAAATAAACTGAGCTACGAAAA 752

RESULT 4

US-10-117-937-84
; Sequence 84, Application US/10117937
; Publication No. US20030220239A1
; GENERAL INFORMATION:
; APPLICANT: CTL IMMUNO THERAPIES CORP.
; APPLICANT: SIMARD, John, J.L.
; APPLICANT: DIAMOND, David, C.
; APPLICANT: LIU, Liping
; APPLICANT: XIE, Zhidong
; TITLE OF INVENTION: EPITOPE SEQUENCES
; FILE REFERENCE: CTILIM.027A
; CURRENT APPLICATION NUMBER: US/10/117,937
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/282,211
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/337,017
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/363,210
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 602
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84
; LENGTH: 752
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-117-937-84

Query Match 100.0%; Score 752; DB 13; Length 752;
Best Local Similarity 100.0%; Pred. No. 6.3e-201;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTCTGCGGCGCTGACCTTCTCTGAGAGCGGGGAGAGGCTCGGAGCCATGACAG 60
DB 1 ATCTCTGCGGCGCTGACCTTCTCTGAGAGCGGGGAGAGGCTCGGAGCCATGACAG 60
QY 61 CCGAAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAG 120
DB 61 CCGAAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAG 120
QY 121 TTCTGTATGCGCCAGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGG 180
DB 121 TTCTGTATGCGCCAGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGG 180
QY 181 GAGGTCCCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAG 240
DB 181 GAGGTCCCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAG 240
QY 241 GTCCGATGCGCGCGGCTTCAAGGCTGAATGATGCTGATGCGCGCGGCGGCGGCG 300
DB 241 GTCCGATGCGCGCGGCTTCAAGGCTGAATGATGCTGATGCGCGCGGCGGCGGCGG 300
QY 301 CCGAAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAG 360
DB 301 CCGAAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAGGCGGGGAGAG 360
QY 361 AGCTGCGCGGAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 420
DB 361 AGCTGCGCGGAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 420
QY 421 TGAAGAGTTCACTGTGTCCGCAACATCTGATTAATGATGCTGATGCTGATGCTGAT 480
DB 421 TGAAGAGTTCACTGTGTCCGCAACATCTGATTAATGATGCTGATGCTGATGCTGAT 480
QY 481 GCCAAGTCAAGTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 540
DB 481 GCCAAGTCAAGTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 540

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 318-3168
 TELEFAX: (212) 752-5958
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 752 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-09-751-798-1

Query Match 100.0%; Score 752; DB 9; Length 752;
 Best Local Similarity 100.0%; Pred. No. 6,3e-201;
 Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATCTCTGTGGGCGCTGACCTTCTCTCTGAGAGCCGGGCAAGAGGCTCCGAGCCATGCAAG 60
Db 1 ATCTCTGTGGGCGCTGACCTTCTCTCTGAGAGCCGGGCAAGAGGCTCCGAGCCATGCAAG 60
QY 61 CCGAAGGCGGGGCAACAGGGGGTTGACAGGGCGATGCTGATGSCCCAGAGAGGCGCTTGGA 120
Db 61 CCGAAGGCGGGGCAACAGGGGGTTGACAGGGCGATGCTGATGSCCCAGAGAGGCGCTTGGA 120
QY 121 TTCCTGATGSCCCAGAGGGGCAATGCTGAGCGCCCAAGAGAGAGCGGATGCAAGGCGGCA 180
Db 121 TTCCTGATGSCCCAGAGGGGCAATGCTGAGCGCCCAAGAGAGAGCGGATGCAAGGCGGCA 180
QY 181 GAGGTCCCGGGGCGAGGGGGCAAGAGGCTCCGGGGCGGGAGAGAGGCGCCCGGGG 240
Db 181 GAGGTCCCGGGGCGAGGGGGCAAGAGGCTCCGGGGCGGGAGAGAGGCGCCCGGGG 240
QY 241 GTCCGATGAGCGGCGCGGCTTCAAGGGCTGATGATGCTGCAAGTCCGGGCGGCGAGGGG 300
Db 241 GTCCGATGAGCGGCGCGGCTTCAAGGGCTGATGATGCTGCAAGTCCGGGCGGCGAGGGG 300
QY 301 CGAGAGAGCGGCGCTGTTGATTTTACTGCGCCATGCTTTCGGAGAACCAATGGAAGCAG 360
Db 301 CGAGAGAGCGGCGCTGTTGATTTTACTGCGCCATGCTTTCGGAGAACCAATGGAAGCAG 360
QY 361 AGCTGGCCGCGAGAGAGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 361 AGCTGGCCGCGAGAGAGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 TGAAGAGATTCACTGTGTCCGGCAACATCTGATCTGATCTGATCTGATCTGATCTGAT 480
Db 421 TGAAGAGATTCACTGTGTCCGGCAACATCTGATCTGATCTGATCTGATCTGATCTGAT 480
QY 481 GCCAAGTGCAGCTCTCATAGCTCTGCTGCTGAGAGAGCTTCCCTGTTGATGATGATCA 540
Db 481 GCCAAGTGCAGCTCTCATAGCTCTGCTGCTGAGAGAGCTTCCCTGTTGATGATGATCA 540
QY 541 CGCAGTGTCTTTCGCGCGGTGTTTGGCTCAAGCTTCCCTCAAGGAGAGAGAGAGAGAG 600
Db 541 CGCAGTGTCTTTCGCGCGGTGTTTGGCTCAAGCTTCCCTCAAGGAGAGAGAGAGAGAG 600
QY 601 AGCCTGGCGCCCTTCTAGATGATGCTCTCCCTAGAGGAATGATCCAGAGAGAGAG 660
Db 601 AGCCTGGCGCCCTTCTAGATGATGCTCTCCCTAGAGGAATGATCCAGAGAGAGAGAG 660
QY 661 GCCAGTTCAATTGTGGGGGCTGATGTTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 661 GCCAGTTCAATTGTGGGGGCTGATGTTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 TTCTGTAGAAATATAAATGAGCTACGAAAAA 752
Db 721 TTCTGTAGAAATATAAATGAGCTACGAAAAA 752

```

RESULT 2
 US-09-849-602-15
 ; Sequence 15, Application US/09849602
 ; Publication No. US20030105834A1
 ; GENERAL INFORMATION:

APPLICANT: Scanlan, Matthew J.
 APPLICANT: Old, Lloyd J.
 APPLICANT: Stockert, Elisabeth
 APPLICANT: Chen, Yao-Tseng
 TITLE OF INVENTION: Colon Cancer Antigen Panel
 FILE REFERENCE: L0461/7105 (US)
 CURRENT APPLICATION NUMBER: US/09/849,602
 NUMBER OF SEQ ID NOS: 30
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 15
 LENGTH: 752
 TYPE: DNA
 ORGANISM: Homo sapien
 US-09-849-602-15

Query Match 100.0%; Score 752; DB 13; Length 752;
 Best Local Similarity 100.0%; Pred. No. 6,3e-201;
 Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATCTCTGTGGGCGCTGACCTTCTCTCTGAGAGCCGGGCAAGAGGCTCCGAGCCATGCAAG 60
Db 1 ATCTCTGTGGGCGCTGACCTTCTCTCTGAGAGCCGGGCAAGAGGCTCCGAGCCATGCAAG 60
QY 61 CCGAAGGCGGGGCAACAGGGGGTTGACAGGGCGATGCTGATGSCCCAGAGAGGCGCTTGGA 120
Db 61 CCGAAGGCGGGGCAACAGGGGGTTGACAGGGCGATGCTGATGSCCCAGAGAGGCGCTTGGA 120
QY 121 TTCCTGATGSCCCAGAGGGGCAATGCTGAGCGCCCAAGAGAGAGCGGATGCAAGGCGGCA 180
Db 121 TTCCTGATGSCCCAGAGGGGCAATGCTGAGCGCCCAAGAGAGAGCGGATGCAAGGCGGCA 180
QY 181 GAGGTCCCGGGGCGAGGGGGCAAGAGGCTCCGGGGCGGGAGAGAGGCGCCCGGGG 240
Db 181 GAGGTCCCGGGGCGAGGGGGCAAGAGGCTCCGGGGCGGGAGAGAGGCGCCCGGGG 240
QY 241 GTCCGATGAGCGGCGCGGCTTCAAGGGCTGATGATGCTGCAAGTCCGGGCGGCGAGGGG 300
Db 241 GTCCGATGAGCGGCGCGGCTTCAAGGGCTGATGATGCTGCAAGTCCGGGCGGCGAGGGG 300
QY 301 CGAGAGAGCGGCGCTGTTGATTTTACTGCGCCATGCTTTCGGAGAACCAATGGAAGCAG 360
Db 301 CGAGAGAGCGGCGCTGTTGATTTTACTGCGCCATGCTTTCGGAGAACCAATGGAAGCAG 360
QY 361 AGCTGGCCGCGAGAGAGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 361 AGCTGGCCGCGAGAGAGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 TGAAGAGATTCACTGTGTCCGGCAACATCTGATCTGATCTGATCTGATCTGATCTGAT 480
Db 421 TGAAGAGATTCACTGTGTCCGGCAACATCTGATCTGATCTGATCTGATCTGATCTGAT 480
QY 481 GCCAAGTGCAGCTCTCATAGCTCTGCTGCTGAGAGAGCTTCCCTGTTGATGATGATGATCA 540
Db 481 GCCAAGTGCAGCTCTCATAGCTCTGCTGCTGAGAGAGCTTCCCTGTTGATGATGATGATCA 540
QY 541 CGCAGTGTCTTTCGCGCGGTGTTTGGCTCAAGCTTCCCTCAAGGAGAGAGAGAGAGAG 600
Db 541 CGCAGTGTCTTTCGCGCGGTGTTTGGCTCAAGCTTCCCTCAAGGAGAGAGAGAGAGAG 600
QY 601 AGCCTGGCGCCCTTCTAGATGATGCTCTCCCTAGAGGAATGATCCAGAGAGAGAGAG 660
Db 601 AGCCTGGCGCCCTTCTAGATGATGCTCTCCCTAGAGGAATGATCCAGAGAGAGAGAG 660
QY 661 GCCAGTTCAATTGTGGGGGCTGATGTTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 661 GCCAGTTCAATTGTGGGGGCTGATGTTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 TTCTGTAGAAATATAAATGAGCTACGAAAAA 752
Db 721 TTCTGTAGAAATATAAATGAGCTACGAAAAA 752

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2004, 12:44:04 : Search time 349 Seconds

(without alignments)
7937.226 Million cell updates/sec

Title: US-10-023-182-1

Perfect score: 752
Sequence: 1 ATCTCGTGGGCGCTGACCT.....TAAACTGACCTACGAAAA 752

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2449703 seqs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	752	100.0	752	9	US-09-751-798-1
2	752	100.0	752	13	US-09-849-602-15
3	752	100.0	752	13	US-10-364-614-15
4	752	100.0	752	13	US-10-117-937-84
5	752	100.0	752	14	US-10-023-182-1
6	743	99.6	752	15	US-10-207-655-70
7	743	98.8	752	15	US-10-295-027-385
8	679.6	90.4	750	15	US-10-146-473-29
9	540.2	71.8	1191	9	US-09-821-883-28
10	434.2	57.7	608	12	US-10-295-027-387
11	413.4	55.0	993	15	US-10-157-031-87
12	127.6	17.0	916	15	US-10-125-237-47
13	127.6	17.0	916	15	US-10-105-891-47
14	117.4	15.6	599	10	US-09-880-107-3826
15	117.4	15.6	599	10	US-09-854-531-176

RESULT 1	US-09-751-798-1	Sequence 1, Appl1				
16	117.4	15.6	599	10	US-09-954-531-385	Sequence 385, App
17	111.6	14.8	429	14	US-10-079-623-228	Sequence 228, App
18	92.4	12.3	400	12	US-10-242-555A-10384	Sequence 10384, A
19	65	8.6	2561	10	US-09-976-740-48	Sequence 48, Appl
20	65	8.6	2561	12	US-10-616-187-48	Sequence 48, Appl
21	65	8.6	2561	14	US-10-023-529-48	Sequence 48, Appl
22	65	8.6	2561	14	US-10-023-523-48	Sequence 48, Appl
23	60	8.0	60	13	US-09-908-975-8697	Sequence 8697, App
24	58.4	7.8	551	13	US-10-029-386-13333	Sequence 13333, A
25	54	7.2	9025608	15	US-10-156-761-1	Sequence 1, Appl1
26	53.2	7.1	154746	13	US-09-827-688-8	Sequence 8, Appl1
27	53.2	7.1	154746	13	US-09-827-688-8	Sequence 8, Appl1
28	53	7.0	3163	12	US-10-292-798-1513	Sequence 1513, App
29	53	7.0	3163	13	US-10-017-161-1857	Sequence 1857, App
30	52.6	7.0	956	13	US-10-027-632-11508	Sequence 31508, A
31	52.6	7.0	956	14	US-10-027-632-11508	Sequence 31508, A
32	52.6	7.0	3824	12	US-10-159-563-107	Sequence 307, App
33	52.6	7.0	3824	13	US-10-116-275-197	Sequence 197, App
34	51.6	6.9	1117	12	US-10-292-798-1141	Sequence 1141, App
35	51.6	6.9	1117	13	US-10-017-161-1403	Sequence 1403, App
36	50.8	6.8	152	13	US-10-029-386-17033	Sequence 17033, A
37	50.4	6.7	3133	12	US-10-292-798-1191	Sequence 1191, App
38	50.4	6.7	3133	13	US-10-017-161-1483	Sequence 1483, App
39	50	6.6	12733	15	US-10-032-393-47	Sequence 47, Appl
40	50	6.6	12733	15	US-10-032-393-47	Sequence 47, Appl
41	49.6	6.6	795	15	US-10-156-761-1729	Sequence 1729, App
42	49.6	6.6	834	15	US-10-156-761-1113	Sequence 1113, App
43	49.6	6.6	1926	13	US-10-294-604-3	Sequence 3, Appl1
44	49.6	6.6	5117	13	US-10-120-988-438	Sequence 438, Appl
45	49.6	6.6	8705	15	US-10-291-230-14	Sequence 14, Appl

US-09-751-798-1
; Sequence 1, Appl1
; Patent No. US20020010321A1
; GENERAL INFORMATION:
; APPLICANT: Stockert, Elisabeth; Jager, Elke;
; APPLICANT: Chen, Yao-tseng; Seaman, Matthew;
; APPLICANT: Knuth, Alexander; Old, Lloyd J.
; TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer
; TITLE OF INVENTION: Associated Proteins, Uses Thereof,
; TITLE OF INVENTION: Truncated Forms of NY-ESO-1, and HLA
; TITLE OF INVENTION: Binding Peptides Derived Therefrom
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski, L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/751,798
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/062,422
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/725,182
; FILING DATE: October 3, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. US20020010321A1man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5466.3

Sequence 780 BP; 131 A; 235 C; 269 G; 145 T; 0 other;

Query Match 98.8%; Score 743; DB 25; Length 780;

Best Local Similarity 100.0%; Pred. No. 5,6e-152; Matches 743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 3 CCTCGTGGGACCTGACCTTCTCTCTGAGAGCCCGGAGAGGCTCCGAGCCATGAGGCC 62
DB 38 CCTCGTGGGACCTGACCTTCTCTCTGAGAGCCCGGAGAGGCTCCGAGCCATGAGGCC 97
QY 63 GAAGGCCGGGGGACAGGGGGGTTGACGGGGGATGCTGATGGCCAGAGGCCCTGGCATT 122
DB 98 GAAGGCCGGGGGACAGGGGGGTTGACGGGGGATGCTGATGGCCAGAGGCCCTGGCATT 157
QY 123 CCTGATGGCCAGGGGGGCAATGCTGGGGGCCAGAGAGGCGGGTGCCACGGGGCGGAGA 182
DB 158 CCTGATGGCCAGGGGGGCAATGCTGGGGGCCAGAGAGGCGGGTGCCACGGGGCGGAGA 217
QY 183 GGTCCCCGGGGGGGAGAGGGGCGAGAGGGGCGGGGAGAGGGGCGGGGGGT 242
DB 218 GGTCCCCGGGGGGGAGAGGGGCGAGAGGGGCGGGGAGAGGGGCGGGGGGT 277
QY 243 CCGCATGGCGGGCGGGCTTCAAGGGCTGAATGATGCTGCAATGGCGGGCCAGGGGGCGG 302
DB 278 CCGCATGGCGGGCGGGCTTCAAGGGCTGAATGATGCTGCAATGGCGGGCCAGGGGGCGG 337
QY 303 GAGAGCCGGCTGCTGAGTTCTAATCTCCGCAATGCTTTGGGAGCACCCATGGAAGCAAG 362
DB 338 GAGAGCCGGCTGCTGAGTTCTAATCTCCGCAATGCTTTGGGAGCACCCATGGAAGCAAG 397
QY 363 CTGGCCCGGAGAGGCTGGGCCAGAGATGCCACCGCTTCCGTTGCGCAGGGGTGCTCTG 422
DB 398 CTGGCCCGGAGAGGCTGGGCCAGAGATGCCACCGCTTCCGTTGCGCAGGGGTGCTCTG 457
QY 423 AAGGAGTTCACTGTGTCCGGCAACATATGACTATCCGACTGCTGCAAGACCAACGCG 482
DB 458 AAGGAGTTCACTGTGTCCGGCAACATATGACTATCCGACTGCTGCAAGACCAACGCG 517
QY 483 CAAGTCAGAGCTTCATCATGCTCCGTCTCCAGAGCTTCCCTGTGATGTGATCAG 542
DB 518 CAAGTCAGAGCTTCATCATGCTCCGTCTCCAGAGCTTCCCTGTGATGTGATCAG 577
QY 543 CAGTGTCTTCCGCGGTGTTTGGCTCAGGCTCCCTCAGGGGAGAGGGGCTAAGCCAG 602
DB 578 CAGTGTCTTCCGCGGTGTTTGGCTCAGGCTCCCTCAGGGGAGAGGGGCTAAGCCAG 637
QY 603 CCGGGGGCCCTTCCAGGTATGCTCCCTAGGGAATGTCCTCAGACAGAGTGC 662
DB 638 CCGGGGGCCCTTCCAGGTATGCTCCCTAGGGAATGTCCTCAGACAGAGTGC 697
QY 663 CAGTTCATTGGGGGGCTGATGTTTGTCTGCTGAGAGAGAGCGGCTTACATGTTTGT 722
DB 698 CAGTTCATTGGGGGGCTGATGTTTGTCTGCTGAGAGAGAGCGGCTTACATGTTTGT 757
QY 723 CTGTAGAAAATAAACTAGACTA 745
DB 758 CTGTAGAAAATAAACTAGACTA 780
```

Search completed: February 13, 2004, 12:43:57
Job time : 286 secs

KW class II restricted T cell epitope; MHC-II epitope; cancer antigen;
 KM NY ESO-1 protein; CD4+ T lymphocyte; human leucocyte antigen; HLA;
 KM tumour-specific humoral-mediated immunity; cancer; cytostatic;
 KM immunotherapy; ds.
 XX Homo sapiens.
 OS
 FT Key Location/Qualifiers
 FT 89..631
 FT /*tag= a
 FT CDS /product= "Human NY ESO-1 protein"
 FT
 PD WO200155393-A2.
 PD 02-AUG-2001.
 PD 26-JAN-2001; 2001WO-US02765.
 PD 28-JAN-2000; 2000US-0179004.
 PD 29-SEP-2000; 2000US-0237107.
 PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Wang R, Rosenberg SA, Zeng G;
 PI WPI; 2001-496851/54.
 DR P-PSDB; AAE07714.
 XX
 PT New NY-ESO cancer peptide or MHC class II restricted T cell epitopes,
 PT useful as immunogen and vaccine for inhibiting cancer in a mammal or as
 PT protection from metastasis.
 PT
 PS Claim 1; Fig 1; 134pp; English.
 XX
 CC The invention relates to the identification and isolation of major
 CC histocompatibility (MHC) class II restricted T cell epitope (MHC-II
 CC epitope) derived from the cancer antigen, NY ESO-1. The MHC-II epitopes
 CC from NY ESO-1 are recognised by CD4+ T lymphocytes in an human leucocyte
 CC antigen (HLA) class II restricted manner, in particular HLA-DR or HLA-DP
 CC restricted. The products of the gene are promising candidates for
 CC immunotherapeutic strategies for the prevention, treatment and diagnosis
 CC of patients with cancer. The cancer epitopes are useful as immunogen and
 CC vaccine to inhibit or to prevent cancer in a mammal by eliciting CD4+ T
 CC lymphocytes resulting in protection of the recipient from development of
 CC cancer and protection from metastasis, or by inhibiting the growth of
 CC cells expressing the NY-ESO-1 gene product. The cancer peptides are also
 CC useful as diagnostic agent to detect the presence of cancer, to enhance
 CC the generation of antibody and/or CD8+ T cell responses against any
 CC given target antigen and/or happen and to induce tumour-specific
 CC humoral-mediated immunity against cancer. The present sequence is
 CC a DNA encoding human NY ESO-1 protein.
 CC Note: The present sequence is also shown in sequence listing of the
 CC specification, but lacks a nucleotide base at its 3' end.
 CC
 XX
 SQ Sequence 806 BP; 155 A; 237 C; 269 G; 145 T; 0 other;

Query Match 99.5%; Score 748.4; DB 22; Length 806;
 Best Local Similarity 99.9%; Pred. No. 3.8e-153;
 Matches 749; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCTCGTGGCCCTGACCTTCTCTGAGAGCCGGGCAAGGCTCCGAGCCATGACGCC 62
 DB 38 CCTCGTGGCCCTGACCTTCTCTGAGAGCCGGGCAAGGCTCCGAGCCATGACGCC 97
 QY 63 GAAGCCGGGGGCAAGAGGGTTTGAAGGCGATGCTGATGGCCAGAGGCCCTTGACATT 122
 DB 98 GAAGCCGGGGGCAAGAGGGTTTGAAGGCGATGCTGATGGCCAGAGGCCCTTGACATT 157
 QY 123 CCTGATGGCCCAAGGGGCAATGCTGGCGGCCAGAGAGGGCGGTGCTCAACGGGCGGCA 182
 DB 158 CCTGATGGCCCAAGGGGCAATGCTGGCGGCCAGAGAGGGCGGTGCTCAACGGGCGGCA 217
 QY 183 GGTCCCGGGGGCGCAAGGGGCAAGAGGGCTCGGGGCGGGAGAGAGCGCCCGCGGGGT 242

DB 218 GGTCCCGGGGGCGCAAGGGGCAAGAGGGCTCGGGGCGGGAGAGAGCGCCCGCGGGGT 277
 QY 243 CCGCATGGCGGGCGGGCTTCAAGGGCTGAATGGATGCTGACAGATGCGGGGCGAGGGGCGG 302
 DB 278 CCGCATGGCGGGCGGGCTTCAAGGGCTGAATGGATGCTGACAGATGCGGGGCGAGGGGCGG 337
 QY 303 GAGAGCCCGCTGCTTGAATTTCACTCGACATGCTTTTGGAGACACCCATGAAACAGAG 362
 DB 338 GAGAGCCCGCTGCTTGAATTTCACTCGACATGCTTTTGGAGACACCCATGAAACAGAG 397
 QY 363 CTGGCCCGGAGAGGCTTGGCCCAAGATGCCACACCGCTTCCCGGCGAGGGGGCTTCTG 422
 DB 398 CTGGCCCGGAGAGGCTTGGCCCAAGATGCCACACCGCTTCCCGGCGAGGGGGCTTCTG 457
 QY 423 AAGAGTTCACTGTCTCGGCAACATAGTACTATCCGACTGACTGTGACAGACACCGC 482
 DB 458 AAGAGTTCACTGTCTCGGCAACATAGTACTATCCGACTGACTGTGACAGACACCGC 517
 QY 483 CAATGAGCTCTCATAGCTCTCTGCTCCAGACAGCTTTCCCTGTGATGATGATCAG 542
 DB 518 CAATGAGCTCTCATAGCTCTCTGCTCCAGACAGCTTTCCCTGTGATGATGATCAG 577
 QY 543 CAGTGCTTCTGCGCGGTTTGTGCTGAGCTCCCTCAGGCGAGAGCGCTAAGCCGAG 602
 DB 578 CAGTGCTTCTGCGCGGTTTGTGCTGAGCTCCCTCAGGCGAGAGCGCTAAGCCGAG 637
 QY 603 CTTGGGCGCCCTTCTAGGTATGCTCTCTCCCTTGAAGGATGTCCAGCACAGGTGCG 662
 DB 638 CTTGGGCGCCCTTCTAGGTATGCTCTCTCCCTTGAAGGATGTCCAGCACAGGTGCG 697
 QY 663 CAGTTATGTTGGGGGCGGCTGATGTTGTGCGTGAAGAGAGAGCGCTTACATGTTGTT 722
 DB 698 CAGTTATGTTGGGGGCGGCTGATGTTGTGCGTGAAGAGAGAGCGCTTACATGTTGTT 757
 QY 723 CTGTAGAAATTAATAGTACGACGAAAA 752
 DB 758 CTGTAGAAATTAATAGTACGACGAAAA 787

RESULT 14
 ID ACC51024 standard; cDNA; 780 BP.
 XX
 AC ACC51024;
 DT 12-JUN-2003 (first entry)
 XX
 DE Human bladder cancer associated cDNA sequence SEQ ID NO:138.
 XX
 KW Human; bladder cancer; cytostatic; gene therapy; vaccine; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO200303906-A2.
 PD 16-JAN-2003.
 PD 03-JUL-2002; 2002WO-US21338.
 PD 03-JUL-2001; 2001US-302814P.
 PD 03-AUG-2001; 2001US-310098P.
 PD 08-NOV-2001; 2001US-343705P.
 PD 13-NOV-2001; 2001US-35066P.
 PD 12-APR-2002; 2002US-372246P.
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 PI Mack DH, Aziz N;
 DR WPI; 2003-201532/19.
 DR P-PSDB; ABR48210.
 XX

Db 721 TGTAGAAATTAACAGACTACGAAAA 749

RESULT 12

AA58599

ID AA58599 standard, cDNA, 806 BP.

XX

AC AAX58599;

XX

DT 16-AUG-1999 (first entry)

XX

DE Human cancer antigen NY ESO-1/CAG-3 cDNA.

XX

NY ESO-1/CAG-3 gene; CAG-3 gene; cancer antigen; human; leukaemia;
XX non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;
XX metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;
XX uterine cancer; breast cancer; prostate cancer; ovarian cancer;
XX cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;
XX liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;
XX vaccine; ss.

XX

OS Homo sapiens.

XX

XX

XX Key Location/Qualifiers

FT 89..631

FT CDS /tag= a

FT /note= "ORF1 (claimed in Claim 2), encodes AA505965"

FT CDS 129..305

FT /tag= b

FT /note= "ORF2 (claimed in Claim 3), encodes AA505966"

XX

XX MO9918206-A2.

XX

XX

XX 15-APR-1999.

XX

XX 21-SEP-1998; 98MO-US19609.

XX

XX 08-OCT-1997; 97US-0061428.

XX

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX

XX Rosenberg SA, Wang RF;

XX

XX WPI: 1999-277270/23.

XX

XX P-BSDb; AA505965, AA505966.

XX

XX Cancer antigen NY ESO1/CAG-3

XX

XX Claim 1; Fig 3a; 88pp; English.

XX

XX This is the nucleotide sequence of a cDNA clone that encodes a novel
XX human cancer antigen, termed NY ESO-1/CAG-3 (or CAG-3). The clone
XX was isolated from a melanoma cell line 586ml library. CAG-3 is
XX a new and potent tumour antigen capable of eliciting an antigen-
XX specific immune response by T cells. Cancer peptides encoded by
XX ORF-1 or ORF-2 of the CAG-3 gene, and their variants, (see AA505965-87)
XX are useful as cancer vaccines that protect the recipient from
XX development of cancer. The invention provides vectors that include
XX DNA encoding a CAG-3 cancer peptide, and transfected or transduced
XX host cells. These vectors and host cells can serve as vaccines in
XX which expression of a cancer peptide results in stimulation of tumour
XX antigen specific T cells in an immunised mammal. The invention also
XX provides: a method of diagnosis of cancer or precancer; a transgenic
XX animal; antisense oligonucleotides that inhibit expression of the
XX cancer peptide; or tumour antigen; antibodies reacting with the CAG-3
XX cancer peptide, useful in diagnostic and detection assays; and
XX methods for preventing or inhibiting cancer by administering a
XX cancer peptide, with or without an HLA molecule. The cancer
XX peptides form part of, or are derived from, cancers such as primary
XX or metastatic melanoma, thymoma, lymphoma, sarcoma, lung cancer,
XX liver cancer, leukemia, uterine cancer, cervical cancer, bladder
XX cancer, kidney cancer and adenocarcinomas such as breast, prostate,
XX ovarian, pancreatic and thyroid cancers. Melanoma is treated by

CC	Inducing cancer-specific T cells in vitro for subsequent return to a patient.
XX	Sequence 806 BP; 155 A; 237 C; 269 G; 145 T; 0 other;
SO	Query Match 99.5%; Score 748.4; DB 20; Length 806; Best Local Similarity 99.9%; Pred. No. 3.8e-153; Matches 749; Conservative 0; Mismatches 1; Indels 0; Gaps 0
OY	3 CCTCGGGCCCTGACCTTCTCTGAGAGCCGGGCAAGAGCTCCGAGCCATGCAAGCC 62
Db	38 CCTCGTGGCCCTGACCTTCTCTGAGAGCCGGGCAAGAGCTCCGAGCCATGCAAGCC 97
OY	63 GAAAGCCGGGGCAAGGGGGTTTCAACGGGCGATGCTGATGCGCCAGAGGCCCTGGCAT 122
Db	98 GAAAGCCGGGGCAAGGGGGTTTCAACGGGCGATGCTGATGCGCCAGAGGCCCTGGCAT 157
OY	123 CCTGATGAGCCCAAGGGGGCAATGCTGCGGGCCCAAGAGAGCGGGATGCCAAGGGCGGACA 182
Db	158 CCGATGAGCCCAAGGGGGCAATGCTGCGGGCCCAAGAGAGCGGGATGCCAAGGGCGGACA 217
OY	183 GGTTCCTGGGGGGGCAAGGGGGCAAGAGGCCCTGGGGGGGGGAGAGAGCGCCCGCGGGT 242
Db	218 GGTTCCTGGGGGGGCAAGGGGGCAAGAGGCCCTGGGGGGGGGAGAGAGCGCCCGCGGGT 277
OY	243 CCGCATGAGGGGGCGGGCTTTCAGAGGGCTGATGATGCTGACATGCGGGGCAAGGGGGCGG 302
Db	278 CCGCATGAGGGGGCGGGCTTTCAGAGGGCTGATGATGCTGACATGCGGGGCAAGGGGGCGG 337
OY	303 GAAAGCCGGCTGCTTGAATTTCTACTCGGCCATGCTTTGGCGACACCCATGGAAGAGAG 362
Db	338 GAAAGCCGGCTTGAATTTCTACTCGGCCATGCTTTGGCGACACCCATGGAAGAGAG 397
OY	363 CTGGCCCGCAGAGAGCTGAGCCCAAGAGATGCCACACCGCTTCGCTGCGAGGGTGTTCG 422
Db	398 CTGGCCCGCAGAGAGCTGAGCCCAAGAGATGCCACACCGCTTCGCTGCGAGGGTGTTCG 457
OY	423 AAGGAATTTACATGTGTCCGGGCAATACCTGATTCGGAATGATGCTGTGAGACCAACGGC 482
Db	458 AAGGAATTTACATGTGTCCGGGCAATACCTGATTCGGAATGATGCTGTGAGACCAACGGC 517
OY	483 CAACCTGACAGCTCTCATCAGCTCCGTGCTCCAGCAGAGCTTTCCTGTGATGTGATCAG 542
Db	518 CAACCTGACAGCTCTCATCAGCTCCGTGCTCCAGCAGAGCTTTCCTGTGATGTGATCAG 577
OY	543 CAGTGTCTTTCGCCCGCTGTTTGTGGCTCAGCGCTCCCAAGGGAGAGGGCTTAAGCCGAG 602
Db	578 CAGTGTCTTTCGCCCGCTGTTTGTGGCTCAGCGCTCCCAAGGGAGAGGGCTTAAGCCGAG 637
OY	603 CCTGGGCGCCCTCTCTAGAGTATGCTCTCCCTTAAGGAGATGATCCCAAGAGTGGC 662
Db	638 CCTGGGCGCCCTCTCTAGAGTATGCTCTCCCTTAAGGAGATGATCCCAAGAGTGGC 697
OY	663 CAGTTATTTGTGGGGGCTGATTTGTTGTGCTTGAAGAGAGAGCGCTTACATGTTTGT 722
Db	698 CAGTTATTTGTGGGGGCTGATTTGTTGTGCTTGAAGAGAGAGCGCTTACATGTTTGT 757
OY	723 CTGTAGAAAATAAACTGAGCTACGAAAAA 752
Db	758 CTGTAGAAAATAAACTGAGCTACGAAAAA 787
RESULT 13	
AD14179	AD14179 standard; DNA; 806 BP.
XX	AD14179;
XX	AD14179;
XX	06-NOV-2001 (first entry)
XX	Human NY ESO-1 DNA.
DE	Human; major histocompatibility complex; MHC; vaccine; metastasis;


```

XX Homo sapiens.
OS
XX MO200107917-A1.
XX
XX 01-FEB-2001.
XX
XX 14-JUL-2000; 2000WO-US19220.
XX
XX 23-JUL-1999; 99US-0359503.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX (SLOK-) SLOAN KETTERING INST CANCER RES.
XX (CORR-) CORNELL RES FOUND INC.
XX
XX Jager E, Stockert E, Old LJ, Knuth A, Chen Y, Scanlan M;
XX
XX WPI; 2001-182822/18.
XX
XX P-PSDB; AAB69946.
XX
XX Method useful for determining the status (e.g. progression, regression
XX or stability of the disease) of a cancerous condition, involves
XX determining the levels of NY-ESO-1 specific antibodies in a sample
XX taken from a patient
XX
XX Example 5; Fig 3; 50pp; English.
XX
XX The present sequence encodes human NY-ESO-1. It is provided in a
XX specification relating to a method for determining the status of a
XX cancerous condition in a patient with a tumour that expresses NY-ESO-1.
XX The method comprises assaying a sample taken from the patient for
XX antibodies that specifically bind to the NY-ESO-1 and comparing the
XX value obtained to a prior value obtained from assay of a prior sample
XX taken from the patient. Any difference between the values is indicative
XX of a change in status of the cancerous condition. The method is useful
XX for determining whether a cancerous condition is progressing, regressing
XX or remaining stable, in particular in patients receiving treatment for a
XX melanoma, adenocarcinoma, non-small cell lung carcinoma or bladder
XX carcinoma.
XX
XX Sequence 752 BP; 126 A; 230 C; 256 G; 140 T; 0 other;
XX
Query Match 100.0%; Score 752; DB 22; Length 752;
Best Local Similarity 100.0%; Pred. No. 6.3e-154;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 421 TGAAGAGTTCACTGTGTCCGGCAACATACACTATCCAGTACTGTGCAAGCACCC 480
DB 421 TGAAGAGTTCACTGTGTCCGGCAACATACACTATCCAGTACTGTGCAAGCACCC 480
QY 481 GCCAAGTGAAGTCTCCCATCAGCTCCCTCCGAGAGCTTCCCTGTTATGTGAATCA 540
DB 481 GCCAAGTGAAGTCTCCCATCAGCTCCCTCCGAGAGCTTCCCTGTTATGTGAATCA 540
QY 541 CCAGTGTCTTCTGCGCGGTGTTTGTAGCTAGCCTCCCTCAGAGGAGGCGCTTAAGCCC 600
DB 541 CCAGTGTCTTCTGCGCGGTGTTTGTAGCTAGCCTCCCTCAGAGGAGGCGCTTAAGCCC 600
QY 601 AGCTGGGCGCCCTTCTTGAAGTCAAGCTCCCTCCCTAGAGGAAATGTCGCCAGCAGAGTG 660
DB 601 AGCTGGGCGCCCTTCTTGAAGTCAAGCTCCCTCCCTAGAGGAAATGTCGCCAGCAGAGTG 660
QY 661 GCAGTTCATGTGGGAGGCGCTGATGTTGTGCTGAGAGAGAGAGGAGCTTAATGTTGT 720
DB 661 GCAGTTCATGTGGGAGGCGCTGATGTTGTGCTGAGAGAGAGAGGAGCTTAATGTTGT 720
QY 721 TTCTGTAGAAATTAATACTGAGCTACGAAAAA 752
DB 721 TTCTGTAGAAATTAATACTGAGCTACGAAAAA 752

RESULT 7
AAS02254
ID AAS02254 standard; DNA; 752 BP.
AC AAS02254;
XX
XX 18-JUL-2001 (first entry)
XX
XX Human NY-ESO-1 tumour rejection antigen precursor DNA sequence.
XX
XX NY-ESO-1; human; tumour rejection antigen precursor; SSX-2; MHC Class II;
XX major histocompatibility complex; helper T cell; HLA-DR; cancer; ds;
XX human leukocyte antigen-determining region; disease progression;
XX disease regression; disease onset; body tissue; body fluid; enzyme label;
XX radioactive label; monoclonal antibody.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 54..596
XX /tag= a
XX /product= "Human NY-ESO-1 protein"
XX
XX MO200123560-A2.
XX
XX 05-APR-2001.
XX
XX 26-SEP-2000; 2000WO-US26411.
XX
XX 29-SEP-1999; 99US-0408036.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Tureci O, Sahin U, Pfeundschnuh M;
XX
XX WPI; 2001-266156/27.
XX
XX P-PSDB; AAU01535.
XX
XX Polypeptides binding to major histocompatibility complex class II human
XX leukocyte antigen-determining region molecule having amino acid
XX sequence found in tumor rejection antigen precursor used for
XX stimulating proliferation of helper T cells -
XX
XX Example 2; Fig 3; 62pp; English.
XX
XX The sequence represents a human NY-ESO-1 DNA. NY-ESO-1 is a tumour
XX rejection antigen precursor. NY-ESO-1 and SSX-2 polypeptides, or

```

```

Db      421 TGAAGAGTTCACTGTCGTCGCGACAACTACTGACTTACCACTGACTGCTGACAGACACC 480
QY      481 GCCAATGACAGCTCTCCATCAGCTCCTGTCTCCAGACAGCTTCCCTGTGATGATGATCA 540
Db      481 GCCAATGACAGCTCTCCATCAGCTCCTGTCTCCAGACAGCTTCCCTGTGATGATGATCA 540
QY      541 CGAGAGCTTCTGCGCCGCTGTTTGGCTGACGCTCCCTCAGGAGACAGGCGCTTAAGCCC 600
Db      541 CGAGAGCTTCTGCGCCGCTGTTTGGCTGACGCTCCCTCAGGAGACAGGCGCTTAAGCCC 600
QY      601 AGCCTGAGCCGCTTCTAGATGATGCTCTCCCTCAGGAGATGCTCCAGACAGATG 660
Db      601 AGCCTGAGCCGCTTCTAGATGATGCTCTCCCTCAGGAGATGCTCCAGACAGATG 660
QY      661 GCCAGTTCAATTGCGGAGGCTGATTTGTTGTGCTGAGAGAGAGAGCGCTTAATGTTGT 720
Db      661 GCCAGTTCAATTGCGGAGGCTGATTTGTTGTGCTGAGAGAGAGAGCGCTTAATGTTGT 720
QY      721 TTCTGTAGAAAAATAAACTAGCTACGAAAAA 752
Db      721 TTCTGTAGAAAAATAAACTAGCTACGAAAAA 752

RESULT 5
AAH75118 standard; cDNA; 752 BP.
XX      AAH75118;
XX      13-NOV-2001 (first entry)
XX      Nucleotide sequence of cancer testis tumour antigen NY-ESO-1.
DE      Cancer testis tumour antigen; NY-ESO-1; LAGE-2; human leukocyte antigen;
KM      HLA; HLA binding peptide; major histocompatibility complex; MHC;
KW      tumour; cancer; testis tumour; ss.
OS      Homo sapiens.
XX      OS
XX      Key      Location/Qualifiers
FH      CDS      54..596
FT      /*tag= a
FT      /product= "cancer testis tumour antigen NY-ESO-1"
PN      MO200162917-A1.
XX      PD      30-AUG-2001.
XX      PF      22-JAN-2001; 2001WC-US02126.
XX      PR      22-FEB-2000; 2000US-0510635.
XX      PA      (LUDW-) LUDWIG INST CANCER RES.
XX      PI      Lethe B, Boon-Falleur T;
XX      DR      WPI: 2001-550091/61.
XX      DR      P-PSDB; AAG67164.
XX      PT      Genomic sequences of tumour associated antigen EY-ESO-1 (LAGE-2) useful
XX      PS      for diagnosing testicular tumours -
XX      PS      Example 5; Fig 3; 50pp; English.
XX      CC      The present sequence encodes cancer testis tumour antigen NY-ESO-1
XX      CC      (also called LAGE-2). NY-ESO-1 is a molecule that is processed to at
XX      CC      least one human leukocyte antigen (HLA) binding peptide, which binds
XX      CC      to Class I and Class II major histocompatibility complex (MHC).
XX      CC      NY-ESO-1 is expressed in tumour mRNA and in testis, but not normal
XX      CC      colon, kidney, liver or brain tissue. The presence or level of expression
XX      CC      of NY-ESO-1 may be assayed for the diagnosis of cancer, especially
XX      CC      testis tumours.

```

```

XX      SQ      Sequence 752 BP; 126 A; 230 C; 256 G; 140 T; 0 other;
XX      Query Match      100.0%; Score 752; DB 22; Length 752;
XX      Best Local Similarity 100.0%; Pred. No. 6,3e-154;
XX      Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 ATCCCTGAGGCGCCCTGACCTTCTCTGAGAGCGCGGCAAGAGGCTCCGAGCATGAGG 60
Db      1 ATCCCTGAGGCGCCCTGACCTTCTCTGAGAGCGCGGCAAGAGGCTCCGAGCATGAGG 60
QY      61 CCGAAGCCCGGAGCACAAGGGGGTTCAGCGGGCGATGCTGATGCGCCACAGAGGCGCTTGGCA 120
Db      61 CCGAAGCCCGGAGCACAAGGGGGTTCAGCGGGCGATGCTGATGCGCCACAGAGGCGCTTGGCA 120
QY      121 TTCTGTAGTGGCCCAAGGGGGCAATGCTGGCGCCCAAGAGAGGGGGTCCACAGGGCGGCA 180
Db      121 TTCTGTAGTGGCCCAAGGGGGCAATGCTGGCGCCCAAGAGAGGGGGTCCACAGGGCGGCA 180
QY      181 GAGGTCCCGGGGCGCAGAGGGCGCAGAGAGGCTTCGGGGCCGGAGAGAGCGCCCGGG 240
Db      181 GAGGTCCCGGGGCGCAGAGGGCGCAGAGAGGCTTCGGGGCCGGAGAGAGCGCCCGGG 240
QY      241 GTCCGATGAGGCGCGGCTTCAGGGCTGAATGATGCTGAGATGCGGGGCCAGAGGGGC 300
Db      241 GTCCGATGAGGCGCGGCTTCAGGGCTGAATGATGCTGAGATGCGGGGCCAGAGGGGC 300
QY      301 CGAGAGCGCGCTGCTTGAAGTTCTACCTCGCAGATGCTTTCGAGACACCCATGAGACAG 360
Db      301 CGAGAGCGCGCTGCTTGAAGTTCTACCTCGCAGATGCTTTCGAGACACCCATGAGACAG 360
QY      361 AGCTGGCCCGCAGAGAGCTTGCGCCAGATGCGCCACCGCTTCCCGTCCAGAGGGTCTTC 420
Db      361 AGCTGGCCCGCAGAGAGCTTGCGCCAGATGCGCCACCGCTTCCCGTCCAGAGGGTCTTC 420
QY      421 TGAAGAGTTCACTGTGTCGGGCAACATCTGATATCCGACTGACTGCTGACAGACACC 480
Db      421 TGAAGAGTTCACTGTGTCGGGCAACATCTGATATCCGACTGACTGCTGACAGACACC 480
QY      481 GCCAATGACAGCTCTCCATCAGCTCCTGTCTCCAGACAGCTTCCCTGTGATGATGATCA 540
Db      481 GCCAATGACAGCTCTCCATCAGCTCCTGTCTCCAGACAGCTTCCCTGTGATGATGATCA 540
QY      541 CGAGAGCTTCTGCGCCGCTGTTTGGCTGACGCTCCCTCAGGAGACAGGCGCTTAAGCCC 600
Db      541 CGAGAGCTTCTGCGCCGCTGTTTGGCTGACGCTCCCTCAGGAGACAGGCGCTTAAGCCC 600
QY      601 AGCCTGAGCCGCTTCTAGATGATGCTCTCCCTCAGGAGATGCTCCAGACAGATG 660
Db      601 AGCCTGAGCCGCTTCTAGATGATGCTCTCCCTCAGGAGATGCTCCAGACAGATG 660
QY      661 GCCAGTTCAATTGCGGAGGCTGATTTGTTGTGCTGAGAGAGAGAGCGCTTAATGTTGT 720
Db      661 GCCAGTTCAATTGCGGAGGCTGATTTGTTGTGCTGAGAGAGAGAGCGCTTAATGTTGT 720
QY      721 TTCTGTAGAAAAATAAACTAGCTACGAAAAA 752
Db      721 TTCTGTAGAAAAATAAACTAGCTACGAAAAA 752

RESULT 6
AAF58634 standard; cDNA; 752 BP.
XX      AAF58634;
XX      AC      27-APR-2001 (first entry)
XX      DT      Human NY-ESO-1 nucleotide sequence.
XX      DE      Human NY-ESO-1; HLA; human leukocyte antigen; CTL; cytotoxic T cell;
XX      KW      HLA-A2; HLA-DR53; melanoma; adenocarcinoma; bladder carcinoma;
XX      KW      non-small cell lung carcinoma; tumour status determination; ss.

```

```

Db      61  CCGAAGCGCGGCGACAGGGGGTTGACAGCGCGATGCTGATGCGCCAGAGAGCCCTTGCA 120
Qy      121  TTCTGATGAGCCAGAGGGGCAATGCTGAGCGGCCAGAGAGCGGGTGGCCACGAGCGGCA 180
Db      121  TTCTGATGAGCCAGAGGGGCAATGCTGAGCGGCCAGAGAGCGGGTGGCCACGAGCGGCA 180
Qy      181  GAGGTCCCCCGGCGAGAGGGGCAAGAGGCTTGAGGGGCGGAGAGAGCGCCCGCGGG 240
Db      181  GAGGTCCCCCGGCGAGAGGGGCAAGAGGCTTGAGGGGCGGAGAGAGCGCCCGCGGG 240
Qy      241  GTCCGATGAGCGCGCGGCTTCAAGGGCTGAATGATGCTGAGATGCGGGGCGCAGGGGCG 300
Db      241  GTCCGATGAGCGCGCGGCTTCAAGGGCTGAATGATGCTGAGATGCGGGGCGCAGGGGCG 300
Qy      301  CGAGAGCGCGCTGCTGATGATTTCTCTGCGCATGCTTTTGGCGACACCCATGGAACAG 360
Db      301  CGAGAGCGCGCTGCTGATGATTTCTCTGCGCATGCTTTTGGCGACACCCATGGAACAG 360
Qy      361  AGCTGCGCGAGAGCGCTGAGTTCTACCTGCGCATGCTTTGCGAGCACCCATGGAACAG 360
Db      361  AGCTGCGCGAGAGCGCTGAGTTCTACCTGCGCATGCTTTGCGAGCACCCATGGAACAG 360
Qy      421  TGAAGAGTTCACTGTGTCCGCAACATACTGACTATCCGACTGACTGCTGCAAGCACCC 480
Db      421  TGAAGAGTTCACTGTGTCCGCAACATACTGACTATCCGACTGACTGCTGCAAGCACCC 480
Qy      481  GCGAAGTGAAGTCTTCATCAGCTCCTGCTCCAGAGCTTTCCCTGTTGATGATGATCA 540
Db      481  GCGAAGTGAAGTCTTCATCAGCTCCTGCTCCAGAGCTTTCCCTGTTGATGATGATCA 540
Qy      541  CGCAGTGTCTTCTGCGCGGTGTTTGGCTCAAGCTTCCCTCAAGGCGAGGCGCTAAGCCC 600
Db      541  CGCAGTGTCTTCTGCGCGGTGTTTGGCTCAAGCTTCCCTCAAGGCGAGGCGCTAAGCCC 600
Qy      601  AGCTGCGCGCGCTTCTGAGTCAATGCGCTTCCCTCAAGGCGAGGCGCTAAGGCTG 660
Db      601  AGCTGCGCGCGCTTCTGAGTCAATGCGCTTCCCTCAAGGCGAGGCGCTAAGGCTG 660
Qy      661  GCGAGTTCATTTGCGGGGCTGATTTGTTGCTGCGAGAGAGAGCGCTTACATTTTGT 720
Db      661  GCGAGTTCATTTGCGGGGCTGATTTGTTGCTGCGAGAGAGAGCGCTTACATTTTGT 720
Qy      721  TTCTGAGAAATTAATACTGAGCTACGAAAA 752
Db      721  TTCTGAGAAATTAATACTGAGCTACGAAAA 752

RESULT 4
AAZ38380
ID      AAZ38380 standard; cDNA; 752 BP.
XX
XX      AAZ38380;
AC
XX      15-FEB-2000 (first entry)
DT
XX
DE      Human tumour antigen NY-ESO-1 cDNA.
XX
XX      Cancer; tumour; antigen; MHC; major histocompatibility complex;
XX      T-cell; cytotoxic; helper; stimulation; proliferation; treatment;
XX      diagnosis; prevention; melanoma; breast cancer; ovarian cancer;
XX      prostate cancer; hepatoma; thyroid cancer; bladder cancer; lung cancer;
XX      lymphoma; ds.
OS      Homo sapiens.
XX
XX      Key
XX      CDS
XX      Location/Qualifiers
XX      FT
XX      CDS
XX      /tag= a
XX      /product= "Human tumour antigen NY-ESO-1"
XX
XX      WO953938-A1.
XX
XX      28-OCT-1999.

```

```

XX      24-MAR-1999; 99WO-US06875.
XX
XX      17-APR-1998; 98US-0062422.
XX
XX      02-OCT-1998; 98US-0165546.
XX
XX      (LUDW-) LUDWIG INST CANCER RES.
XX
XX      Stockert E, Jager E, Chen Y, Scanlan M, Alexander K, Old LJ;
XX      Gure A, Rilter G;
XX      WPI; 2000-038483/03.
XX
XX      P-PSDB; AAY52430.
XX
XX      Novel peptides which bind to MHC class I and MHC class II molecules,
XX      useful for therapeutic and diagnostic purposes -
XX
XX      Claim 30; Fig 3; 49pp; English.
XX
XX      This sequence represents cDNA encoding a human tumour antigen NY-ESO-1,
XX      which was isolated from an oesophagus squamous cell cancer cDNA
XX      library. Tissue localisation studies revealed it to be expressed at
XX      high levels in normal ovary and testis but not in normal colon, kidney,
XX      liver, brain, oesophagus and skin. It was expressed in certain tumours
XX      and tumour cell lines with some degree of frequency - these included
XX      melanoma specimens and cell lines, and breast and bladder cancer
XX      specimens, with expression in other tumour types being sporadic.
XX      Peptides derived from NY-ESO-1 are bound by both MHC (major
XX      histocompatibility complex) Class I and Class II molecules for
XX      presentation to T-cells. Peptides AAY52431-52434 bind to Class I HLA-A2
XX      molecules, thereby stimulating proliferation of cytotoxic T-cells, while
XX      peptides AAY52435-52440 bind to Class II HLA-DM3 molecules, stimulating
XX      helper T-cell proliferation. The peptides derived from NY-ESO-1 may be
XX      used in methods and compositions used for the treatment, diagnosis and
XX      prevention of cancers (such as melanoma, breast cancer, prostate cancer,
XX      lung cancer, hepatoma, ovarian cancer, thyroid cancer, bladder cancer,
XX      or lymphoma) and to stimulate the proliferation of T cells.
XX
XX      Sequence 752 BP; 126 A; 230 C; 256 G; 140 T; 0 other;
XX
XX      Query Match 100.0%; Score 752; DB 21; Length 752;
XX      Best Local Similarity 100.0%; Pred. No. 6.3e-154;
XX      Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1  ATCTCGTGGGCGCTTCACTCTGAGAGCGCGGCGAGGCTCCGAGCGCATGCAAG 60
Db      1  ATCTCGTGGGCGCTTCACTCTGAGAGCGCGGCGAGGCTCCGAGCGCATGCAAG 60
Qy      61  CCGAAGCGCGGCGACAGGGGGTTTGAAGCGGCGATGCTGATGCGCCAGAGAGCGCTTGCA 120
Db      61  CCGAAGCGCGGCGACAGGGGGTTTGAAGCGGCGATGCTGATGCGCCAGAGAGCGCTTGCA 120
Qy      121  TTCTGATGAGCCAGAGGGGCAATGCTGAGCGGCCAGAGAGAGGGGTTGCCACGAGCGGCA 180
Db      121  TTCTGATGAGCCAGAGGGGCAATGCTGAGCGGCCAGAGAGAGGGGTTGCCACGAGCGGCA 180
Qy      181  GAGGTCCCCCGGCGAGAGGGGCAAGAGGCTTGAGGGGCGGAGAGAGCGCCCGCGGG 240
Db      181  GAGGTCCCCCGGCGAGAGGGGCAAGAGGCTTGAGGGGCGGAGAGAGCGCCCGCGGG 240
Qy      241  GTCCGATGAGCGCGCGGCTTCAAGGGCTGAATGATGCTTTGCGAGCACCCATGGAACAG 300
Db      241  GTCCGATGAGCGCGCGGCTTCAAGGGCTGAATGATGCTTTGCGAGCACCCATGGAACAG 300
Qy      301  CGAGAGCGCGCTGCTGATGATTTCTCTGCGCATGCTTTTGGCGACACCCATGGAACAG 360
Db      301  CGAGAGCGCGCTGCTGATGATTTCTCTGCGCATGCTTTTGGCGACACCCATGGAACAG 360
Qy      361  AGCTGCGCGAGAGCGCTGAGTTCTACCTGCGCATGCTTTGCGAGCACCCATGGAACAG 420
Db      361  AGCTGCGCGAGAGCGCTGAGTTCTACCTGCGCATGCTTTGCGAGCACCCATGGAACAG 420
Qy      421  TGAAGAGTTCACTGTGTCCGCAACATACTGACTATCCGACTGACTGCTGCAAGCACCC 480

```

CC cancer, and can be utilised as a marker for the targeted delivery of
 CC therapeutic agents to oesophageal cancer cells. It can also be used to
 CC generate diagnostic or therapeutic agents.

XX
 SQ Sequence 752 BP; 126 A; 230 C; 256 G; 140 T; 0 other;

Query Match 100.0%; Score 752; DB 21; Length 752;
 Best Local Similarity 100.0%; Pred. No. 6.3e-154;
 Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATCTCTGAGGAGCCCTGACCTTCTCTCTGAGAGCCGAGGAGGCTCCGAGGAGCCATGACAG 60
DB 1 ATCTCTGAGGAGCCCTGACCTTCTCTCTGAGAGCCGAGGAGGCTCCGAGGAGCCATGACAG 60
QY 61 CCGAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
DB 61 CCGAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
QY 121 TTCTGTATGAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
DB 121 TTCTGTATGAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
QY 181 GAGGTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
DB 181 GAGGTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
QY 241 GTCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
DB 241 GTCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
QY 301 CGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
DB 301 CGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
QY 361 AGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
DB 361 AGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY 421 TGAAGAGGATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 421 TGAAGAGGATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 481 GCCAAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 481 GCCAAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 541 CGCAGTCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
DB 541 CGCAGTCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
QY 601 AGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
DB 601 AGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
QY 661 GCCAGTCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
DB 661 GCCAGTCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
QY 721 TTCTGTAGAAAAATAAAGTCAAGAAAAA 752
DB 721 TTCTGTAGAAAAATAAAGTCAAGAAAAA 752

```

RESULT 3
 ID AAD00152
 AAD00152 standard; cDNA; 752 BP.

XX AAD00152;
 AC
 XX 31-JUL-2000 (first entry)
 DT
 XX Human tumour antigen, NY-ESO-1 cDNA.
 DE
 XX

KW NY-ESO-1; CAMEL; CTL-recognised Antigen on Melanoma; human; cancer; CTL;
 KW cytotoxic T lymphocyte; tumour-associated antigen; LAGE-1; anticancer;
 KW melanoma; immunotherapy; immune response; ss.
 XX
 OS Homo sapiens.

XX Key Location/Qualifiers
 FX 5'UTR 1..53
 FT /tag= a
 FT 54..596
 FT CDS /tag= b
 FT /product= "Human NY-ESO-1 protein"
 FT /note= "Derived from open reading frame (ORF)-1"
 FT 597..752
 FT 3'UTR /tag= c
 FT 1..93
 FT 5'UTR /tag= d
 FT /note= "5' UTR of ORF-2"
 FT CDS /tag= e
 FT /product= "Human NY-ESO-1 short variant protein"
 FT /note= "Derived from open reading frame (ORF)-2"
 FT 271..752
 FT /tag= f
 FT /note= "3' UTR of ORF-2"

PN W0200023564-A1.

PD 27-APR-2000.

PF 15-OCT-1999; 99WO-EP07832.

PR 16-OCT-1998; 98EP-0119583.

PA (BOEH) BOEHRINGER INGELHEIM INT GMBH.
 PA (UHO-) UNIV HOSPITAL LEIDEN.

PI Schrier PI, Aarnoudse CA, Heider K, Klade C;

DR WPI: 2000-339685/29.

DR P-PDB: AAY70862, AAY70863.

PT Tumor-associated antigen useful for cancer immunotherapy is encoded by
 PT the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -
 XX
 PS Example 1; Page 61-62; 73pb; English.

XX The present sequence is the cDNA encoding human NY-ESO-1, a tumour
 CC antigen, identified by screening a cDNA library of an oesophagus
 CC carcinoma. It has two different reading frames and the proteins decoded
 CC contain epitopes of tumour specific T-cells. NY-ESO-1 is expressed in
 CC different tumour types, but not in healthy tissues except in testis. It
 CC also shows homology with the CAMEL (Cytotoxic T lymphocytes (CTL)-
 CC recognised Antigen on Melanoma) cDNA, a tumour-associated antigen. The
 CC tumour-associated antigen displayed on melanoma cells is recognised by
 CC cytotoxic T lymphocytes. This sequence has anticancer activity. CAMEL
 CC tumour antigen and immunogenic peptides derived from it are useful for
 CC cancer immunotherapy. They have the potential to induce an immune
 CC response, by eliciting a CTL response. The DNA molecule is used for the
 CC construction of recombinant or fusion proteins.
 XX

SQ Sequence 752 BP; 126 A; 230 C; 256 G; 140 T; 0 other;

Query Match 100.0%; Score 752; DB 21; Length 752;
 Best Local Similarity 100.0%; Pred. No. 6.3e-154;
 Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATCTCTGAGGAGCCCTGACCTTCTCTCTGAGAGCCGAGGAGGCTCCGAGGAGCCATGACAG 60
DB 1 ATCTCTGAGGAGCCCTGACCTTCTCTCTGAGAGCCGAGGAGGCTCCGAGGAGCCATGACAG 60
QY 61 CCGAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
DB 61 CCGAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120

```


PI Old LJ, Scanlan M;
 XX MPI; 1998-286417/25.
 DR P-PSDB; AAM62584.
 XX
 XX
 PT New isolated cancer associated antigen - is used to develop products
 PT for the diagnosis and treatment of cancers and for monitoring cancer
 PT therapy
 XX
 XX
 PS Claim 1; Fig 3; 49pp; English.

CC The present sequence encodes a cancer associated antigen. The clone
 CC from which the DNA sequence is obtained is designated NY-ESO-1. The
 CC specification described a method for determining regression, progression
 CC of onset of a cancerous condition, comprising monitoring a sample from a
 CC patient with the cancerous condition for a parameter selected from
 CC NY-ESO-1 protein, a peptide derived from NY-ESO-1 protein and cytolytic
 CC T cells specific for the peptide and an MHC molecule with which it
 CC non-covalently complexes. Methods for the treatment of a cancerous
 CC condition are also described. The NY-ESO-1 protein and peptides derived
 CC from it can be used for diagnosis and treatment of cancers and to monitor
 CC the efficacy of a therapeutic regime.
 CC
 XX
 XX

SQ Sequence 752 BP; 126 A; 230 C; 256 G; 140 T; 0 other;

Query Match 100.0%; Score 752; DB 19; Length 752;

Best Local Similarity 100.0%; Pred. No. 6,3e-154;
 Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTCTGTGGGCGCTTACCTTCTCTGTGAGAGCCGGGCGAGAGCTCCGAGCCATGACGAG 60
 DB 1 ATCTCTGTGGGCGCTTACCTTCTCTGTGAGAGCCGGGCGAGAGCTCCGAGCCATGACGAG 60
 QY 61 CCGAAGGCGGGGCGACAGAGGGGTTTCAGAGCGGCGATGCTGATGAGCCCGAGAGGCGCTTGCA 120
 DB 61 CCGAAGGCGGGGCGACAGAGGGGTTTCAGAGCGGCGATGCTGATGAGCCCGAGAGGCGCTTGCA 120
 QY 121 TTCTGTATGAGCCCGAGGGGCGATGCTGAGCGGCGCCGAGAGAGGCGGGTCCACGGGGGCGCA 180
 DB 121 TTCTGTATGAGCCCGAGGGGCGATGCTGAGCGGCGCCGAGAGAGGCGGGTCCACGGGGGCGCA 180
 QY 181 GAGGTCCCGGGGGCGAGAGGGGCGTCCGAGGCGCGGAGAGAGGCGCGCGCGGCG 240
 DB 181 GAGGTCCCGGGGGCGAGAGGGGCGTCCGAGGCGCGGAGAGAGGCGCGCGCGGCG 240
 QY 241 GTCCGATGAGCGGCGGCTTCAGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 DB 241 GTCCGATGAGCGGCGGCTTCAGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 QY 301 CGAGAGAGCGCGCTTGAATGCTTACCTGCGCAAGCTTTCGAGAGAGCCGAGAGAGAGAGAGAGAG 360
 DB 301 CGAGAGAGCGCGCTTGAATGCTTACCTGCGCAAGCTTTCGAGAGAGCCGAGAGAGAGAGAGAGAG 360
 QY 361 AGCTGGCGCGCAGAGAGCTGCGCCAGAGATGCCACCGCTTCCCGTCCGAGAGAGAGAGAGAGAG 420
 DB 361 AGCTGGCGCGCAGAGAGCTGCGCCAGAGATGCCACCGCTTCCCGTCCGAGAGAGAGAGAGAGAG 420
 QY 421 TGAAGAGATGCTGATGCTGCGGAGCAACTACTGACTATCGAGTCTGCTGAGAGAGAGAGAGAGAG 480
 DB 421 TGAAGAGATGCTGATGCTGCGGAGCAACTACTGACTATCGAGTCTGCTGAGAGAGAGAGAGAGAG 480
 QY 481 GCGAAGTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 DB 481 GCGAAGTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 QY 541 GCGAGTCTTCTGCGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 DB 541 GCGAGTCTTCTGCGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 QY 601 AGCTGGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 DB 601 AGCTGGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660

QY 661 GCGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 DB 661 GCGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 721 TTCTGTAGAAAATATACTGAGCTACGAAAAA 752
 DB 721 TTCTGTAGAAAATATACTGAGCTACGAAAAA 752

RESULT 2

AAA61483 ID AAA61483 standard; cDNA; 752 BP.

AC AAA61483;

DT 23-OCT-2000 (first entry)

DE cDNA encoding human oesophageal cancer-associated antigen NY-ESO-1.

KW Oesophageal cancer associated antigen; NY-ESO-1; human;

KW Immunogen; oesophageal carcinoma; melanoma; ovary; testis;

KW transmembrane domain; antibody; diagnostic marker; drug delivery target;

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 54..596

FT /tag= a /product= "Oesophageal cancer-associated antigen,

FT /note= "Nucleotides 54-600 are specifically claimed"

PN US6069233-A.

PD 30-MAY-2000.

PF 26-JAN-1998; 98US-0013150.

PR 03-OCT-1996; 96US-0725381.

PA (SLOK) SLOAN KETTERING INST CANCER RES.

PA (CORR) CORNELL RES FOUND INC.

PA (LUDM-) LUDWIG INST CANCER RES.

PI Scanlan M, Gure AO, Chen Y, Tureci O, Sahin U, Pfeundschn M;

PI Old LJ;

PT MPI; 2000-410880/35.

PT P-PSDB; AAB03154.

PS Claim 1; Fig 3; 9pp; English.

CC This sequence represents cDNA encoding a human oesophageal cancer-
 CC associated antigen, NY-ESO-1. This sequence was isolated from a
 CC cDNA library prepared from a specimen of well-to-moderately
 CC differentiated squamous cell cancer of the oesophagus. Expression
 CC analysis demonstrated that NY-ESO-1 was expressed in oesophageal
 CC carcinoma, certain melanoma cell lines and in normal ovary and testis
 CC tissue, but not in normal colon, kidney, liver or brain tissue. Analysis
 CC of the amino acid sequence of the protein indicates that the protein has
 CC a transmembrane domain, several N-myristoylation sites and
 CC O-phosphorylation sites and that it contains antigenic sequences in the
 CC N-terminal half of the protein. The antigen is useful as an immunogen
 CC when combined with an adjuvant, in both precursor and post-
 CC translationally modified forms, and may be used to generate anti-NY-ESO-1
 CC antibodies. It can also be used as a diagnostic marker for oesophageal

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2004, 01:55:07 ; Search time 284 Seconds
(without alignments)
7147.808 Million cell updates/sec

Title: US-10-023-182-1

Perfect score: 1 ATCCCTGCTGGCCCTGACCT.....TAAACTGACCTACGAAAA 752

Scoring table: IDENTITY NUC

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 510512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N Geneseq 19Jun03:*

1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*
25: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	752	100.0	752	19	AAV38566
2	752	100.0	752	21	AAAG1483
3	752	100.0	752	21	AAAD00152
4	752	100.0	752	21	AAZ38380
5	752	100.0	752	22	AAH75118
6	752	100.0	752	22	AAAF5834
7	752	100.0	752	22	AA502254
8	752	100.0	752	24	ABK84456

9	752	100.0	752	25	ABX96556
10	752	100.0	752	25	ABQ83850
11	749	99.6	755	19	AAV50348
12	748.4	99.5	806	20	AAK58599
13	748.4	99.5	806	22	AAD14179
14	743	98.8	780	25	ACC51024
15	743	98.8	780	25	ABX76232
16	743	98.8	780	25	ABX76423
17	681.6	90.6	767	21	AAD00150
18	680.6	90.5	755	19	AAV50347
19	680.6	90.5	755	24	ABV72284
20	599.2	79.7	679	21	AAD00149
21	540.2	71.8	1191	22	AAD21573
22	527.4	70.1	540	22	AAD14180
23	442.6	58.9	993	21	AAV50346
24	438.6	58.3	1002	19	AAV50346
25	434.2	57.7	608	25	ACC51025
26	434.2	57.7	608	25	ABX76233
27	434.2	57.7	608	25	ABX76424
28	372	49.5	873	22	AAH75122
29	323.4	43.0	1751	22	AAH75123
30	323.4	43.0	2630	22	AAH75124
31	215.8	28.7	5572	22	AAH46407
32	209.4	27.8	211	22	AAD14184
33	181.4	24.1	458	22	AAH75121
34	137.8	18.3	5572	22	AAH46408
35	117.4	15.6	599	24	ABN97331
36	117.4	15.6	599	24	ABL62906
37	117.4	15.6	599	24	ABL63115
38	117.4	15.6	599	24	ABK35550
39	96.8	12.9	10623	24	ABK35950
40	83.4	11.1	217	19	AAV50349
41	75	10.0	90	22	AAD14181
42	65	8.6	90	24	ABK36928
43	65	8.6	2561	22	AAH26500
44	64.4	8.6	90	24	ABK36923
45	64.4	8.6	90	24	ABK36939

ALIGNMENTS

RESULT 1	
AAV38566	
ID	AAV38566 standard; DNA; 752 BP.
XX	
AC	AAV38566;
XX	
DT	17-SEP-1998 (first entry)
XX	
DE	DNA encoding a cancer associated antigen.
XX	
KW	Cancer associated antigen; NY-ESO-1; regression; progression; onset; cancer; treatment; diagnosis; ds.
XX	
OS	Homo sapiens.
XX	
HH	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	/*tag= a
FT	/product= cancer associated antigen
XX	
PN	W09814464-A1.
XX	
PD	09-APR-1998.
XX	
EF	15-SEP-1997; 97WO-US16335.
XX	
PR	03-OCT-1996; 96US-0725182.
XX	
PA	(LUDW-) LUDWIG INST CANCER RES.
XX	
FI	Chen Y, Drifftout JW, Gure A, Jager E, Knuth A;

Human cDNA encoding Human NY-ESO-1 enc Human NY-ESO-1 nuc Human cancer antig Human NY-ESO-1 DNA Human NY-ESO-1 cod Lung cancer-associ Lung cancer-associ Human LAGE-1 splic Human LAGE-1 clone Human LAGE-1 clone Human CTL-recogis Human NY-ESO-1 IC-HB Human NY-ESO-1 cod Human LAGE-1 unspl Human LAGE-1 clone Human bladder canc Lung cancer-associ Lung cancer-associ Lung cancer-associ Cancer testis tumo Cancer testis tumo Nucleotide sequenc Tumour suppressor Human NY-ESO-1 MHC Cancer testis tumo Tumour suppressor Gene #3829 used to Breast cancer rela Human DNA encoding Human DNA encoding Human DNA encoding Human DNA encoding

QY 221 GGGAGAGAGGCGCCCGGAGGTCGCGCATGCGCGCGGCTTCAAGGCTGAATGATGCTG 280
 DB 180 GAGAGAGAGGCGCCCGGAGGTCGCGCATGCGCGCGGCTTTCGCGAGATGAGAGTG 239
 QY 281 CAGATGCGGAGGCGCGAGAGGCGCGCTTCTGAAGTTTCACTTCCGATGCTTT 340
 DB 240 CCGCTGC-GGGCCAGAGAGCGGACAGCGCTCTTCAATTGACATCAAGATGCTTT 298
 QY 341 CGGACACCCCATGAGAGAGAGCTGCGCGGAGAGCTGCGCGGATGCGGATGCGGATG 400
 DB 299 CTGTCGCGCATGAGAGAGAGCTGTCGCGAGATCTTTCGCGGATGCGGATGCTTT 358
 QY 401 TCCGTCGCGAGAGGCTTCTGAAGAGTTCACTGTGTCCGCAACACTACTACTAT 457
 DB 359 CCGCCGACAGAGGCGGTTCTGAAGAGCTTCACTGTGTCCGCAACACTACTACTAT 415
 RESULT 14
 LOCUS B1856831 856 bp mRNA linear EST 10-OCT-2001
 DEFINITION 603387624F1 NIH_MGC_87 Homo sapiens CDNA clone IMAGE:5396573 5',
 mRNA sequence.
 ACCESSION B1856831
 VERSION B1856831 GI:15997578
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 856)
 NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DP
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM12010 row: 0 column: 06
 High quality sequence stop: 850.
 Location/Qualifiers
 1..856
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5396573"
 /tissue_type="mammary adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 87"
 /note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.383 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
 BASE COUNT 149 a 263 c 298 g 145 t 1 others
 ORIGIN
 Query Match 42.3%; Score 317.8; DB 12; Length 856;
 Best Local Similarity 89.3%; Pred. No. 1.7e-62;
 Matches 376; Conservative 0; Mismatches 42; Indels 3; Gaps 3;
 QY 37 GCGAGAGCTCCGAGCATGACAGCGAGCGGCGCAAGGCGGTTGACGCGCGATG 96
 DB 1 GCGAGAGCTCCGAGCATGACAGCGAGCGGCGCAAGGCGGTTGACGCGCGATG 59
 QY 97 CTGATGCGCGAGAGGCGCTTCTGATGCGCCAGGAGGCGCAATGCTGGCGGCGCAG 156
 DB 60 CTGATGCGCGAGAGGCGCTTCTGATGCGCCAGGAGGCGCAATGCTGGCGGCGCAG 118

QY 157 GAGAGCGGAGTCCAGCGGCGCGAGAGTCCCGGCGCGAGGCGCGCAAGGCGCTTCG 216
 DB 119 GAGAGCGGAGTCCAGCGGCGCGAGAGTCCCGGCGCGAGGCGCGCAAGGCGCTTCG 178
 QY 217 GCGCGGAGAGGCGCGCGGCGGAGTCCGAGTGGGCGGCGGCTTCAAGGCTGAATGAT 276
 DB 179 GCGCGGAGAGGCGCGCGGCGGAGTCCGAGTGGGCGGCGGCTTTCGCGAGATGGA 238
 QY 277 GCTGAGATGCGGAGGCGCGGCGCGGAGAGCTTCTGAGTTTCACTTCCGATGC 336
 DB 239 GGTGCGCGGCGCGAGAGGCGCGGAGAGCTTCTGAGTTTCACTTCCGATGC 298
 QY 337 CTTCGCGACACCATGAGAGAGAGCTGCGCGGAGAGCTGAGGCGGATGCGGATG 396
 DB 299 CTTCGCGCGCATGAGAGAGAGCTGTCGCGAGATCTTTCGCGGATGCGGATG 358
 QY 397 CGCTTCCGTCGCGAGAGGCTTCTGAAGAGTTCACTGTGTCCGCAACACTACTACTA 456
 DB 359 CTTCGCGCGCATGAGAGAGAGCTGTCGCGAGATCTTTCGCGGATGCGGATG 417
 QY 457 T 457
 DB 418 T 418
 RESULT 15
 LOCUS B1859796 764 bp mRNA linear EST 10-OCT-2001
 DEFINITION 603386106F1 NIH_MGC_87 Homo sapiens CDNA clone IMAGE:539961 5',
 mRNA sequence.
 ACCESSION B1859796
 VERSION B1859796 GI:16000543
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 764)
 NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DP
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM12006 row: 1 column: 02
 High quality sequence stop: 526.
 Location/Qualifiers
 1..764
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5394961"
 /tissue_type="mammary adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 87"
 /note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.383 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
 BASE COUNT 100 a 234 c 283 g 147 t
 ORIGIN
 Query Match 41.7%; Score 313.8; DB 12; Length 764;
 Best Local Similarity 89.5%; Pred. No. 1.3e-61;
 Matches 416; Conservative 0; Mismatches 42; Indels 7; Gaps 7;

RESULT 12	BM797218	377 bp	mba	linear	EST 05-MAR-2002
LOCUS	K-EST0080221.S22SNUI6n1				
DEFINITION	5', mRNA sequence.				
ACCESSION	BM797218				
VERSION	GI:19145450				
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 377)				
AUTHORS	Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R., Oh, K.O., Cheong, J.E., Sohn, H.Y., Kim, J.W., Park, H.S., Kim, S. and Kim, Y.S.				
TITLE	21C Frontier Korean EST Project 2001				
JOURNAL	Unpublished				
COMMENT	Contact: Kim YS Genome Research Center Korea Research Institute of Bioscience & Biotechnology 52, Eoseun-dong Yuseong-gu, Daejeon 305-353, South Korea Tel.: +82-42-860-4470 Fax: +82-42-860-4409 Email: yongsung@mail.krrib.re.kr Plate: 73 row: E column: 05 High quality sequence stop: 377. Location/Qualifiers 1..377 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="S22SNUI6n1-73-E05" /sex="F" /tissue_type="Ascltea" /cell_type="Lymphoblast-like" /cell_line="SNU-16" /lab_host="DH10B" /clone_id="S22SNUI6n1" /note="Organ: Stomach; Vector: pT7T3-Pac; Site_1: EcoRI; Site_2: NotI; The S22SNUI6 library was contributed by the Soares Laboratory and it was constructed as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1995), Genome Research 6(9): 791-806. RNA was prepared from harvested cells of SNU-16 culture. SNU-16 cell was obtained from Korean Cell Line Bank (KCLB). SNU-16 was established from ascitic fluids of Korean patients by Park J.G. et al. (1990), Cancer Res 50: 2773-2780."				
BASE COUNT	55 a 121 c 149 g 52 t				
ORIGIN					
Query Match	42.9%; Score 322.8; DB 12; Length 377;				
Best Local Similarity	92.6%; Pred. No. 1e-63;				
Matches 339; Conservative 0; Mismatches 27; Indels 0; Gaps 0;					
1	ATCCCTGCGGGGCCCGACCTTCTCTGTAGAGCGCGGGGAGAGGCTCCGAGCATATGACG 60				
12	ATCTCTGCGGGGCCCGACCTTCTCTGTAGAGCGCGGGGAGAGGCTCCGAGCATATGACG 71				
61	CCGAGGCGCGGGGCGACAGGCGGGTTTGACAGCGCGATGCTGATGCGCCAGAGGCGCTTGAC 120				
72	CCGAGGCGCGGGGCGACAGGCGGGTTTGACAGCGCGATGCTGATGCGCCAGAGGCGCTTGAC 131				
121	TTCTGATGCGCCAGAGGCGCAATGCTTGAGCGGCCCGAGAGAGCGGGGTGCTCAACGAGCGGCA 180				
132	TTCTGATGCGCCAGAGGCGCAATGCTTGAGCGGCCCGAGAGAGCGGGGTGCTCAACGAGCGGCA 191				
181	GAGGTCCCGGGGCGCGAGGGGCGAGAGGCGCTCGGGGCGGGAGAGAGCGCCCGCGCG 240				
192	CAGGTCCCGGGGCGCGAGGGGCGAGAGGCGCTCGGGGCGGGAGAGAGAGCGCCCGCGCG 251				

QY	241	GTCCGATGGCGGGCGGCTTCAGAGGGGTGATGATCTGTAGATGCGGGCCAGGGGGC	300
Db	252	GTCCGATGGCGGGCGGCTTCGCGAGATGAAGTCCCTGGGGGCGACAGAGC	311
QY	301	CGAGAGCCCGCTGCTTGAAGTTACTCTGCCCATGCTTTTCGGACACCCATGAGACAG	360
Db	312	CGAGACACCCCGCTTCAGTTGCATCAATCAGATGCTTCTGTGCGCCCATGAGAGCGG	371
QY	361	AGCTGG 366	
Db	372	AGCTGG 377	
RESULT 13			
LOCUS	BG575764	770 bp	mRNA linear EST 10-APR-2001
DEFINITION	602597257F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4705839 5', mRNA sequence.		
ACCESSION	BG575764		
VERSION	BG575764.1	GI:13583417	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryotic Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	1 (bases 1 to 770)		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabds-remail.nih.gov		
	Tissue Procurement: DCTD/DTF		
	CDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov		
	Plate: LHAM0573 row: b column: 16		
FEATURES	High quality sequence stop: 713.		
Source	Location/Qualifiers		
	1..770		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4705839"		
	/tissue_type="mammary adenocarcinoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/clone_id="NIH_MGC_87"		
	/note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."		
BASE COUNT	135 a 231 c 270 g 134 t		
ORIGIN			
Query Match	42.9%	Score 322.6;	DB 10; Length 770;
Best Local Similarity	89.0%;	Pred. No. 1.3e-63;	
Matches 371; Conservative	0;	Mismatches 44;	Indels 2; Gaps 2;
QY	41	AGGCTCCGAGCCTATGACGCTCCAAAGCCCGGGGCGACAGGGGTTCACAGGCGCATGCTGA	100
Db	1	AGGCTCCGAGCCTATGACGCTCCAAAGCCAGTGACAGGGGGTTCACAGGCGCATGCTGA	60
QY	101	TGCGCCAGAGAGCCCTGAGATTCTGATGGCCAGGGGCGCATGCTGGCGCCAGAGAGA	160
Db	61	TGCGCCAGAGAGCCCTGAGATTCTGATGGCCA-ggggcaatctggcgccagagaga	119
QY	161	GGCGGATGCCAGCGGGCGGCAAGATCCCCGGGGCGACAGGGCGACAGGGCTTGGGGCC	220
Db	120	GGCGGATGCCAGCGGGCGGCAAGATCCCCGGGGCGACAGGGCGACAGGGCTTGGGGCC	179

RESULT 7
BM684237/c 502 bp mRNA linear EST 27-FEB-2002
LOCUS
DEFINITION UI-E-BJ1-aj1-m-18-0-UI.s1 UI-E-BJ1 Homo sapiens cDNA clone
ACCESSION BM684237
VERSION BM684237.1 GI:18994133
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 502)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PubMed 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Forward
POLA=yes
Location/Qualifiers
1..502
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-BJ1-aj1-m-18-0-UI"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-BJ1"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-BJ1 is a subtracted cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are: fetal eyes, AGAATCAAGA
; lens, CGATTAGCGA; eye anterior segment, AATGCCGAT;
optic nerve, CCAATAGG; retina, CCGCG; Retina Foveal and
Macular, GTCC; RPE and Choroid, ACCCTA. This library was
created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI).
TAG_LIB=UI-E-BJ1
TAG_TISSUE=human retina
TAG_SEQ=CCGCG"

BASE COUNT 111 a 140 c 153 g 97 t 1 others
ORIGIN

Query Match 56.6%; Score 425.6; DB 12; Length 502;
Best Local Similarity 91.8%; Pired. No. 2,9e-87;
Matches 449; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 264 GGGCTGATGATGATGCTGAGATGGCGGGGCCAGAGGGGGCCGAGAGCGGCTGCTGAGTTC 323
DB 502 GCGCAGATGAGAGAGTGCCCTTCGCGGGGCCAGAGAGCCCGACAGCCGCTTCCTCACTTG 443
QY 324 TACCTCGCCATGACCTTTTGGCAGACACCCATGGAGACAGAGCTGAGCCGCGAGAGGCTGAGCC 383
DB 442 CACATCAGAGAGCTTTTCTGCTGCGCCATGAGAGCGAGCTGCTCCGACAGATCTGTTC 383
QY 384 CAGATGCCCCACAGCTTCCCTGTCAGAGGGTGTCTTGAAGAGAGTCACTGTGCGGC 443
DB 382 CCGAGATGCGGACCTCTTCCCGACAGAGGGTGTCTTGAAGAGAGTCACTGTGCGGC 323
QY 444 AACATCTGACTATCTGCACTGACTGCTGACAGACACCGCCAACTGACAGCTTCCATCAGC 503
DB 322 AACCTACTGTTTATCCGACTACTGCTGACAGACACCGCCAACTGACAGCTTCCATCAGC 263
QY 504 TCTGTCTTCAGCAGACTTCCCTGTTGATGTGATCAGCAGAGTCTTTCGCCGTTC 563
DB 262 TCTGTCTTCAGCAGACTTCCCTGTTGATGTGATCAGCAGAGTCTTTCGCCGTTC 203
QY 564 TTGGCTCAGGCTTCCCTCAGAGGAGGAGGCTTAAGCCAGGCTGAGCCCTTCTAGGTC 623
DB 202 TTGGCTCAGGCTTCCCTCAGAGGAGGAGGCTTAAGCCAGGCTGAGCCCTTCTAGGTC 143
QY 624 ATGCTCTCTCCCTTGAAGAGATGTCTCCAGCAGAGTGGCCAGTTGTTGGGGGCTTGA 683
DB 142 ATGCTCTCTCCCTTGAAGAGATGTCTCCAGCAGAGTGGCCAGTTGTTGGGGGCTTGA 83
QY 684 TTGTTTGTGCTGAGAGAGAGGAGGCTTACATGTTGTTGTAAGAAATAAAGTGAAC 743
DB 82 TTGTTTGTGCTGAGAGAGAGGAGGCTTACATGTTGTTGTAAGAAATAAAGTGAAC 23
QY 744 TACGAAAA 752
DB 22 TACGAAAA 14

RESULT 8
BE385880/c 852 bp mRNA linear EST 21-JUL-2000
LOCUS
DEFINITION 601275948F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3617123 5',
mRNA sequence.
ACCESSION BE385880
VERSION BE385880.1 GI:9331245
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 852)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-rc@mail.nih.gov
Tissue Procurement: ATCC/DC/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LINC283 row: o column: 12
High quality sequence start: 14
High quality sequence stop: 520.
Location/Qualifiers
1..852
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3617123"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"

QY 170 CACGGGCGGCAAGATGCCCGGGGCGCAGAGGGGCGACGAGGCTCGGGGCGGGAGAGG 229
 DB 596 CACGGGCGGCAAGATGCCCGGGGCGCAGAGGGGCGACGAGGCTCGGGGCGGGAGAGG 537
 QY 230 CGCGCCCGGGGGGCGCATGCGCGCGCGCTTCAGGGGCTGAATGATGCTGCAGATGCGG 289
 DB 536 CGCGCCCGGGGGGCGCATGCGCGCGCGCTTCAGGGGCTGAATGATGCTGCAGATGCGG 477
 QY 290 GCGCAGAGGGGCGGAGAGCGCGCTGCTTGAATTCATCTCGCATGCCCTTTCGCGACACC 349
 DB 476 GCGCAGAGGGGCGGAGAGCGCGCTGCTTGAATTCATCTCGCATGCCCTTTCGCGACACC 417
 QY 350 CATGAGAGGAGAGCTGCGCGCGCGCGCTGCGCGAGATGCGCGACGCTTCGCGACACC 409
 DB 416 CATGAGAGGAGAGCTGCGCGCGCGCGCTGCGCGAGATGCGCGACGCTTCGCGACACC 357
 QY 410 AGGGGCTGCTTGAAGAGATTCAGTGTCTCGCGCAACATATGACTATCCGACTGCTGC 469
 DB 356 AGGGGCTGCTTGAAGAGATTCAGTGTCTCGCGCAACATATGACTATCCGACTGCTGC 297
 QY 470 TGCAGACCCCGCCCACTGAGCTCTCCATGAGCTCTCTCCAGAGCTTTCCTGCTT 529
 DB 296 TGCAGACCCCGCCCACTGAGCTCTCCATGAGCTCTCTCCAGAGCTTTCCTGCTT 237
 QY 530 GATGTGATGACGCGAGTGTCTGCGCGCTGCTTGTGCTCAGAGCTCTCCAGAGGAG 589
 DB 236 GATGTGATGACGCGAGTGTCTGCGCGCTGCTTGTGCTCAGAGCTCTCCAGAGGAG 177
 QY 590 GCGCTAAGCCGAGCTGCGCGCGCGCTTCTGATGATGCTCTCTCCAGAGGAGTGTGC 649
 DB 176 GCGCTAAGCCGAGCTGCGCGCGCGCTTCTGATGATGCTCTCTCCAGAGGAGTGTGC 117
 QY 650 CAGCAGAGAGTGGCAGTGTCTGCTGCGCGCGCTGATGCTTGTGCTGAGAGAGAGCGCT 709
 DB 116 CAGCAGAGAGTGGCAGTGTCTGCTGCGCGCGCTGATGCTTGTGCTGAGAGAGAGCGCT 57
 QY 710 TACATGTTGTTTCTGTAGAAAATAAACTGAGCTACGAAAA 752
 DB 56 TACATGTTGTTTCTGTAGAAAATAAACTGAGCTACGAAAA 14
 RESULT 6
 BM719406 529 bp mRNA linear EST 01-MAR-2002
 LOCUS UI-E-EJ1-aj1-m-18-0-UI.r1 UI-E-EJ1 Homo sapiens cDNA clone
 DEFINITION UI-E-EJ1-aj1-m-18-0-UI 5', mRNA sequence.
 ACCESSION BM719406
 VERSION BM719406.1 GI:19038052
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 529)
 AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBR, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: Bento-Soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).

FEATURES Seq primer: M13 Reverse
 Location/Qualifiers
 source 1..529
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-EJ1-m-18-0-UI"
 /tissue_type="fetal eyes, lens, eye anterior segment,
 optic nerve, retina, Retina Foveal and Macular, RPE and
 Choroid"
 /dev_stage="fetal and adult"
 /lab_host="DH10B (Life Technologies) ('11 phage resistant')"
 /note="Organ: eye; Vector: pRT73-Pac (pharmacia) with a
 modified polylinker; Site_1: EcoR I; Site_2: Not I;
 UI-E-EJ1 is a subcloned cDNA library constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pRT73-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tags for this library are: fetal eyes, AGAATCAAGA
 ; lens, CGATTAGCGA; eye anterior segment, ATGCGCAT;
 optic nerve, CCATTAGTG; retina, CCGCG; Retina Foveal and
 Macular, GTCC; RPE and Choroid, ACCGA. This library was
 created for the program Gene Discovery in the Visual
 System, supported by National Eye Institute (NEI)."
 BASE COUNT 79 a 172 c 186 g 92 t
 ORIGIN

Query Match 61.4%; Score 461.8; DB 12; Length 529;
 Best Local Similarity 92.1%; Pred. No. 1.5e-95;
 Matches 487; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 42 GCTCCGAGGATGACAGGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 101
 DB 1 GCTCCGAGGATGACAGGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
 QY 102 GGGCCAGAGGCG 161
 DB 61 GGGCCAGAGGCG 120
 QY 162 GCGGAGTGCAGGCGCGCGAGAGTCCCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 221
 DB 121 GCGGAGTGCAGGCGCGCGAGAGTCCCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
 QY 222 GGAGAGGCG 281
 DB 181 AGAGAGGCG 240
 QY 282 AGATGCGGAGGCG 341
 DB 241 CCGTGGGAGGCGCGAGAGGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
 QY 342 GCGACACCATGAGAGGAGAGTGGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 401
 DB 301 TCGTGGCGCGAGAGGAGAGTGGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 QY 402 CCGGCGAGAGGCG 461
 DB 361 CCGGCGAGAGGCG 420
 QY 462 CTGACTGCTGAGACACCGCGCAACTGAGTCTTCATGAGCTCTGCTTCACAGCTT 521
 DB 421 CTGACTGCTGAGACACCGCGCAACTGAGTCTTCATGAGCTCTGCTTCACAGCTT 480
 QY 522 TCCCTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 570
 DB 481 TCCCTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 529

Oy		217	GAGCGGAGAGAGAGCCGCCCGGGGGTCCGACATGAGCGGCAGCCTTCACAGGCTGAATGAT	276
Dd		246	GCGCGGAGAGAGGCGCCCCCGGGGTCCGCAATGCGCGCGCGCGCTTCAGAGGCTGAATGAT	305
Oy		277	GCTGAGATGCGGGGCGAAGGGGCGGAGAGCGCGCTGTGATTATCCTCGCATGCG	336
Dd		306	GCTGAGATGCGGGGCGAAGGGGCGGAGAGCGCGCTGTGATTATCCTCGCATGCG	365
Oy		337	CTTTGGGACAACCCATNGAAGAAGACTGGCCCCGAGAGCGCTGGGCCAAGATGGCCAC	396
Dd		366	CTTTGGGACAACCCATNGAAGAAGACTGGCCCCGAGAGCGCTGGGCCAAGATGGCCAC	424
Oy		397	CGCTTCCCCTGCGCAGGCGGTCTTCTGAAGAAGATTCACTGTGTCCGCAACATACTACTTA	455
Dd		425	CGCTTCCCCTGCGCAGGCGGTCTTCTGAAGAAGATTCACTGTGTCCGCAACATACTACTTA	484
Oy		457	TCCGACTGACTGCTGTGACACCACCGCAACTGACGCTCTCCATCMAGCTCTCTGTCCAGC	516
Dd		485	TCCGACTGACTGCTGTGACACCACCGCAACTGACGCTCTCCATCMAGCTCTCTGTCCAGC	544
Oy		517	AGCTTCCCTGTTGATGTGATCAAGCAGTGCTTTCTGCGCCTGTTTTGGCTCAGCCTC	576
Dd		545	AGCTTCCCTGTTGATGTGATCAAGCAGTGCTTTCTGCGCCTGTTTTGGCTCAGCCTC	603
Oy		577	CCTCAGGCGAGAGCGCTTAGCTCAGCTGCGCGCCCTTCTTAGGTCAATGCTCCTCCCG	636
Dd		604	CCTCAGGCGAGAGCGCTTAGCTCAGCTGCGCGCCCTTCTTAGGTCAATGCTCCTCCCG	663
Oy		637	TAGGAAATGTGTCCAGCACAGATGGCGCATGTATGTGGGGGCGCATGTTGTCGCGT	696
Dd		664	TAGGAAATGTGTCCAGCACAGATGGCGCATGTATGTGGGGGCGCATGTTGTCGCGT	723
Oy		697	GAGAGGAGCGGCTTACATGTTGTTCTCTGTAGAAAAATAAACGTGACGTACGAAAA	752
Dd		724	GAGAGGAGCGGCTTACATGTTGTTCTCTGTAGAAAAATAAACGTGACGTACGAAAA	779
RESULT 3				
AL553953/c			1201 bp	mRNA linear EST 31-MAY-2001
LOCUS				
DEFINITION			AL553953 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA	
ACCESSION			clone CS0D1079YI02 3-PRIME, mRNA sequence.	
VERSION			AL553953	
KEYWORDS			AL553953.2 GI:31275766	
SOURCE			EST.	
ORGANISM			Homo sapiens (human)	
JOURNAL			Homo sapiens	
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS			Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.	
TITLE			1 (bases 1 to 1201)	
COMMENT			Li,W.B., Gruber,C., Tessee,J. and Polayee,D. Full-length cDNA libraries and normalization Unpublished On Feb 15, 2001 this sequence version replaced gi:12894269, Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seq@genoscope.cns.fr Web : www.genoscope.cns.fr Library was constructed by life technologies, a division of Invitrogen. This sequence belongs to sequence cluster 10526.r For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS0D1079BE01NP1&cluster=10526.r. Contact : Feng Jiang Email : fjiang@life tech.com URL : http://fulllength.invitrogen.com/ invitrogen Corporation 1600 Paradise Avenue Genoscope sequence ID : CS0D1079BE01NP1. Location/Qualifiers 1..1201 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0D1079YI02" /tissue_type="PLACENTA COT 25-NORMALIZED"	
FEATURES				
SOURCE				

```

/clone_11b="Homo sapiens PLACENTA COT 25-NORMALIZED"
/ncore="1st strand cDNA was primed with a NotI-0:190 (dr)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      225 a      302 c      323 g      249 t      102 others
ORIGIN

```

Query Match	87.0%	Score 654.6;	DB 9;	Length 1201;
Best Local Similarity	98.0%;	Pred. No. 1.1e-139;		
Matches 686;	Conservative 5;	Mismatches 6;	Indels 3;	Gaps 3;

QY	37	CCAGAGGCTCCGAGGCCATGAGGCGGAGGCGGAGGCGACAGGGGGCTTCAGCGGGCATGTG	96
Db	698	GCGAGGCTTCGAGAGCCATGCA-GCCGAGGCGGGGACAGGGGGTTGCAGCGGGCATGTG	640
QY	97	CTGATGCGCCAGAGGCGCTTGACATTCCTGATGCGCCAGGGGCGCATGCTGGCGGGCCAG	156
Db	639	CTGATGCGCCAGAGGCGCTTGACATTCCTGATGCGCCAGGGGCGCATGCTGGCGGGCCAG	580
QY	157	GAGAGCGGGGTGCCAGCGGCGCGCGAGAGGTCCCGGGGCGCGAGGGGCAACAGAGGCGCTGG	216
Db	579	GAGAGCGGGGTGCCAGCGGCGCGAGAGGTCCCGGGGCGCGAGGGGCAACAGAGGCGCTGG	520
QY	217	GCGCGGAGAGAGGCGCGCCCGCGGGTCCGACATGCGCGCGCGGCTTCAGGGCTGAATGAT	276
Db	519	GCGCGGAGAGAGGCGCGCCCGCGGGGT-CGACATGCGCGCGGCTTCAGGGCTGAATGAT	461
QY	277	GCTGACATGTCGGGGCCAGGGGGCGGAGAGCGCTGCTTGAGTTCTACTCTGCCATGC	336
Db	460	GCTGACATGTCGGGGCCAGGGGGCGGAGAGCGCTGCTTGAGTTCTACTCTGCCATGC	401
QY	337	CTTTCGCAACCCATGGAAGCAGAGCTGAGCCCGAGAGGCTGGGCCAGAGATGGCCCCAC	396
Db	400	CTTTCGCAACCCATGGAAGCAGAGCTGAGCCCGAGAGGCTGGGCCAGAGATGGCCCCAC	341
QY	397	CGCTTCCCGTCCAGGGGTCTTCTGAAGAGTTCACTGTGTCCGGCAACATCTGACTA	456
Db	340	CGCTTCCCGTCCAGGGGTCTTCTGAAGAGTTCACTGTGTCCGGCAACATCTGACTA	281
QY	457	TCCGACTACTGTGCAGACCAACGCGCCCAAGTCAGCTCCATCAGTCCGCTGTCCAGC	516
Db	280	TCCGACTACTGTGCAGACCAACGCGCCCAAGTCAGCTCCATCAGTCCGCTGTCCAGC	221
QY	517	AGCTTCCCTGTATGATGTGATCAACGAGTCTTTCTGCGCGTGTGTCAGGCTC	576
Db	220	AGCTTCCCTGTATGATGTGATCAACGAGTCTTTCTGCGCGTGTGTCAGGCTC	161
QY	577	CCTCAGGCGAAGGGCTTAAGCCACGCGCTGGGCGCCCTTCTTAAGTCAATGCTCCCTCCC	636
Db	160	CCTCAGGCGAAGGGCTTAAGCCACGCGCTGGGCGCCCTTCTTAAGTCAATGCTCCCTCCC	101
QY	637	TAGGGAATGTGCCAGCAACGAGTGGCCAGTTCAATGTGGGGGCCGATGTGTTGTCT- GCT	695
Db	100	TAGGGAATGTGCCAGCAACGAGTGGCCAGTTCAATGTGGGGRCTGATGTGTTGTCTCANN	41
QY	636	GAGGAGGACGCGCTTACATGTTTGTCTTGAGAAAAATA 735	
Db	40	GAGGAGGACGCGCTTACATGTTTGTGNNNDATAGAAAAATA 1	

RESULT 4	BE410697	646 bp	EST 21-JUL-2000
LOCUS	BE410697		
DEFINITION	60130076331 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3635511 5', mRNA sequence.		
ACCESSION	BE410697		
VERSION	BE410697.1		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Bukarjovla; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		

Sequence: 507-580, >GC-rich#low_complexity
Seq primer: M13 FORWARD
POLYA=yes.

FEATURES

SOURCE

Location/Qualifiers

1..726

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-H-E21-Bbf-J-09-0-UT"

/tissue_type="Chondrosarcoma Grade II"

/dev_stage="Adult"

/lab_host="MDH10B (Life Technologies)"

/clone_lib="NCI CGAP Ch2"

/note="Organ: Left Pelvis; Vector: pUT73-Pac (Pharmacia)

with a modified polylinker; Site 1: EcoR I; Site 2: Not I;

NCI CGAP Ch2 is a normalized cDNA library containing the

following tissue(s): Chondrosarcoma Grade II. The library

was constructed according to Bonaldo, Lennon and Soares,

Genome Research, 6:791-806, 1996. First strand cDNA

synthesis was primed with an oligo-dT primer containing a

Not I site. Double stranded cDNA was ligated to an EcoR I

adaptor, digested with Not I, and cloned directionally

into pUT73-Pac vector. The oligonucleotide used to prime

the synthesis of first-strand cDNA contains a library tag

sequence that is located between the Not I site and the

(dT)18 tail. The sequence tag for this library is

TAGACAGCT

TAG-TISUB-grade-2

TAG_SEQ-ATCTAATATG

243 c 217 g 135 t

BASE COUNT

131 a 243 c 217 g 135 t

ORIGIN

Query Match 92.4%; Score 694.6; DB 13; Length 726;

Best Local Similarity 99.2%; Pred. No. 6.5e-149;

Matches 709; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

38 CAGAGGCTCCGAGCCATGACAGGCCGAGAGCCGAGGCGTTCGACGCGCCGATGC 97
726 CAGAGGCTCCGAGCCATGACAGGCCGAGAGCCGAGGCGTTCGACGCGCCGATGC 669
98 TGATGCGCCAGAGAGCCCTGCGATTCGATGAGGCCAGAGGCGCAATGCTGCGGCCAGG 157
668 TGATGCGCCAGAGAGCCCTGCGATTCGATGAGGCCAGAGGCGCAATGCTGCGGCCAGG 609
158 AGAGCGGGGTGCGACGAGCGCGAGAGGTCCCGGGGCGCAGAGGCGCAAGAGGCGCTCGGG 217
608 AGAGCGGGGTGCGACGAGCGCGAGAGGTCCCGGGGCGCAGAGGCGCAAGAGGCGCTCGGG 549
218 GCGGAGAGAGAGCGCCCGCGGGGTCCGATGAGCGCGCGGCTTCAGAGGCGTGAATGATG 277
548 GCGGAGAGAGAGCGCCCGCGGGGTCCGATGAGCGCGCGGCTTCAGAGGCGTGAATGATG 489
278 CTGACAGATCGAGAGCGCGCGAGAGCCGCTGCTTGAATCTTACCTCGCCATGCC 337
488 CTGACAGATCGAGAGCGCGCGAGAGCCGCTGCTTGAATCTTACCTCGCCATGCC 429
338 TTTCCGAGACCCATGAAAGCAGAGTGGCCCGCAGAGAGCTTGGCCCAAGATGCCCAAC 397
428 TTTCCGAGACCCATGAAAGCAGAGTGGCCCGCAGAGAGCTTGGCCCAAGATGCCCAAC 369
398 GCTTCCCGTCCAGAGGCGCTCTGTAAGAGATTCACTGTCTCCGCAACATCACTACTAT 457
368 GCTTCCCGTCCAGAGGCGCTCTGTAAGAGATTCACTGTCTCCGCAACATCACTACTAT 309
458 CCGACTGACTGCTGACACACCGCCAACTGCACTCTTCATCAAGCTCTGTCTCAGCA 517
308 CCGACTGACTGCTGACACACCGCCAACTGCACTCTTCATCAAGCTCTGTCTCAGCA 249
518 GCTTCCCGTCCAGAGGCGCTCTGTAAGAGATTCACTGTCTCCGCAACATCACTACTAT 577
248 GCTTCCCGTCCAGAGGCGCTCTGTAAGAGATTCACTGTCTCCGCAACATCACTACTAT 189

QY 578 CTCAGAGGAGAGAGGCGCTAAGCCAGCCCTGAGGCGCCCTTCTCAGTCAAGTCACTCTCCCT 637
DB 188 CTCAGAGGAGAGAGGCGCTAAGCCAGCCCTGAGGCGCCCTTCTCAGTCAAGTCACTCTCCCT 129
QY 638 AGGAGATGCTCCAGACAGAGTGGCCAGTTCATTTGAGGAGGCGCTGATTTGTCGCTGG 697
DB 128 AGGAGATGCTCCAGACAGAGTGGCCAGTTCATTTGAGGAGGCGCTGATTTGTCGCTGG 69
QY 698 AGGAGAGGCGCTTCAAGTGTGTTCTGTGTAAGAAATATAACTGAGTACGAAAAA 752
DB 68 AGGAGAGGCGCTTCAAGTGTGTTCTGTGTAAGAAATATAACTGAGTACGAAAAA 14

RESULT 2
AL553954 1201 bp mRNA linear EST 31-MAY-2003
LOCUS AL553954 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1079102 5-PRIME, mRNA sequence.
ACCESSION AL553954
VERSION AL553954.2 GI:31275767
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 15, 2001 this sequence version replaced gi:12894271.
Contact: Genoscope - Centre National de Sequencage
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10526.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1079B010P1&cluster=10526.r. Contact:
Feng Liang Email: fliang@lifech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID: CS0D1079B010P1.

FEATURES

SOURCE

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0D1079102"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

Bases of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 221 a 318 c 306 g 244 t 112 others

ORIGIN

Query Match 89.3%; Score 671.4; DB 9; Length 1201;
Best Local Similarity 98.6%; Pred. No. 1.5e-143;

Matches 706; Conservative 2; Mismatches 4; Indels 4; Gaps 3;

QY 37 GCAAGAGCTCCGAGCCATGACAGGCCGAGAGCCGAGGCGTTCGAGCGGCGGATG 96
DB 68 GCAAGAGCTCCGAGCCATGACAGGCCGAGAGCCGAGGCGTTCGAGCGGCGGATG 125
QY 97 CTGATGCGCCAGAGAGGCGCTGCGATTCGATGAGCGCCAGAGGCGCAATGCTGCGGCGCCAG 156
DB 126 CTGATGCGCCAGAGAGGCGCTGCGATTCGATGAGCGCCAGAGGCGCAATGCTGCGGCGCCAG 185
QY 157 GAGAGAGCGAGTGCACAGGCGCGAGAGGTCCCGGGGCGCAGAGGCGAGCAAGGCGCTCG 216
DB 186 GAGAGAGCGAGTGCACAGGCGCGAGAGGTCCCGGGGCGCAGAGGCGAGCAAGGCGCTCG 245

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

CM nucleic - nucleic search, using SW model

Run on: February 13, 2004, 09:58:53 ; Search time 2093 Seconds
(without alignments)
8732.425 Million cell updates/sec

Title: US-10-023-182-1

Perfect score: 752
Sequence: 1 ATCTCGGCGGCCCTGACCT.....TAAACTGAGCTACGAAAA 752

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_esbba:*
2: em_eschum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_escpl:*
7: em_escro:*
8: em_hcc:*
9: gb_esc1:*
10: gb_esc2:*
11: gb_hcc:*
12: gb_esc3:*
13: gb_esc4:*
14: gb_esc5:*
15: em_escfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_pbg:*
27: em_gss_vt:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	694.6	92.4	726	13	B0575599 UI-H-EZ1-
2	671.4	89.3	1201	9	AL553954 AL553954
3	654.6	87.0	1201	9	AL553953 AL553953
4	553	73.5	646	10	BE410697 BE410697

Result No.	Score	Query Match	Length	ID	Description
5	513.2	68.2	596	13	B0629220 UI-H-FLO-
6	461.8	61.4	529	12	BM719406 UI-E-EJ1-
7	425.6	56.6	502	12	BM68437 UI-E-EJ1-
8	376	50.0	852	10	BE385880 601275948
9	355.8	47.3	359	9	AI218223 qh23e03.x
10	348.8	46.4	458	13	BE279602 BX279602
11	323.8	43.1	866	12	BI861547 BI861547
12	322.8	42.9	377	10	BM797218 K-EST0080
13	322.6	42.9	770	10	BC575764 602597257
14	317.8	42.3	856	12	BI856831 BI856831
15	313.8	41.7	764	12	BI859796 BI859796
16	309.8	41.2	689	10	BE274123 BE274123
17	308.4	41.0	566	10	BE390259 BE390259
18	302.4	40.2	383	10	BE410952 BE410952
19	291.6	38.8	564	14	CA424840 CA424840
20	273.6	36.4	739	10	BE391232 BE391232
21	271	36.0	281	9	AA634317 AA634317
22	270.2	35.9	559	10	BE276194 BE276194
23	264.4	35.2	511	10	BE387918 BE387918
24	242.2	32.2	800	10	BE408892 BE408892
25	221	29.4	674	10	BE388562 BE388562
26	217	28.9	639	10	BE387002 BE387002
27	137.6	18.3	707	14	CB462814 CB462814
28	135.4	18.0	544	12	BM973372 BM973372
29	133.2	17.7	534	9	AV663843 AV663843
30	132.4	17.6	635	14	CB991050 CB991050
31	131.6	17.5	1048	13	BQ073801 BQ073801
32	130.4	17.3	736	13	BQ604903 MI-P-CPI-
33	129.6	17.2	590	11	BC019012 Homo sapi
34	129.2	17.2	559	9	AM157381 au88d01.x
35	128.8	17.1	694	14	CA412926 UI-H-EZO-
36	127.6	17.0	631	13	BQ649305 AGENCOURT
37	127.6	17.0	639	10	BQ386671 602454931
38	127.6	17.0	657	13	BQ687170 AGENCOURT
39	127	16.9	528	13	BQ632235 UI-H-FEL-
40	126	16.8	520	14	CA455564 AGENCOURT
41	122.6	16.3	693	10	BE727845 601564226
42	121.6	16.2	592	13	BQ604749 MI-P-CPI-
43	119.4	15.9	1156	12	BM918216 BM918216
44	118.8	15.8	714	9	AV703867 AV703867
45	118.4	15.7	451	14	W73156 zd53b12.s1

ALIGNMENTS

RESULT 1
B0575599/c 726 bp mRNA linear EST 19-JUN-2002
LOCUS
DEFINITION
UI-H-EZ1-bbf-j-09-0-UI.s1 NCI CGAP Ch2 Homo sapiens cDNA clone
UI-H-EZ1-bbf-j-09-0-UI 3', mRNA sequence.

ACCESSION
B0575599
VERSION
B0575599.1 GI:21478916

KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)

REFERENCE
1 (bases 1 to 726)
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS
TITLE
Tumor Gene Index

JOURNAL
CONTACT: Robert Strausberg, Ph.D.
Unpublished
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tissue Procurement: Dr. Steven Gitelis/ Rush Presbyterian, Dept. of
Orthopedics

CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu

The following repetitive elements were found in this cDNA

Db 301 AGAGCCGCTGCTTGAAGTTCTACCTGCGCATGCTTTCGCGACACCCATGAGAGCAGAC 360

QY 364 TGGCCCGCAGAGCCTGAGCCAGAGATGCCCAACCGCTTCCGTCAGAGGGTGTCTTGA 423

Db 361 TGGCCCGCAGAGCCTGAGCCAGAGATGCCCAACCGCTTCCGTCAGAGGGTGTCTTGA 420

QY 424 AGAGTTCACTGTGTGCGGAGCATATACATACATACATACATACATACATACATACAT 483

Db 421 AGAGTTCACTGTGTGCGGAGCATATACATACATACATACATACATACATACATACAT 480

QY 484 AACTGACGCTTCCTCATACAGCTCTCTGTCTCAGAGCTTCCCTGTGATGTGATCAGC 543

Db 481 AACTGACGCTTCCTCATACAGCTCTCTGTCTCAGAGCTTCCCTGTGATGTGATCAGC 540

QY 544 AGTGTCTTTCGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 603

Db 541 AGTGTCTTTCGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600

QY 604 CTGCGCCGCTTCTCTAGTCTAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 663

Db 601 CTGCGCCGCTTCTCTAGTCTAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660

QY 664 AGTTCATTTGGGGGCTGATTTGTTGCTGCTGAGAGAGAGGCTTATGTTTGTTC 723

Db 661 AGTTCATTTGGGGGCTGATTTGTTGCTGCTGAGAGAGAGGCTTATGTTTGTTC 720

QY 724 TGTAGAAATTAATACTGAGCTACGAAAAA 752

Db 721 TGTAGAAATTAATACTGAGCTACGAAAAA 749

RESULT 15

LOCUS HSA003149 755 bp mRNA linear PRI 23-JUN-1998

DEFINITION Homo sapiens mRNA for NY-ESO-1 protein.

ACCESSION AJ003149

VERSION AJ003149.1 GI:3255990

KEYWORDS NY-ESO-1 gene.

SOURCE Homo sapiens (human)

ORGANISM Hukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1

AUTHORS Lethe, B., Lucas, S., Michaux, J., De Smet, C., Godelaine, D.,

TITLE Serrano, A., De Plaen, E. and Boon, T.

JOURNAL LAGE-1, a new gene with tumor specificity

MEDLINE Int. J. Cancer 76 (6), 903-908 (1998)

PUBMED 98289662

REFERENCE 2 (bases 1 to 755)

AUTHORS Lethe, B.G.

TITLE Direct Submision

JOURNAL Submitted (02-JAN-1998) Lethe B.G., Brussels Branch, Ludwig

Brussels, Belgium

location/Qualifiers

1..755

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="X"

/map="q"

/clone="LL-1.2"

/cell_line="LB373-MEL"

/cell_type="melanoma"

/clone_1db="pCDNAL/Amp"

/notes="clone 3"

1..755

/gene="NY-ESO-1"

51..593

/gene="NY-ESO-1"

/codon_start=1

/product="NY-ESO-1 protein"

gene

CDS

/protein_id="CAA05908.1"

/db_xref="GI:3255991"

/db_xref="SWISS-PROT:P78358"

/translation="MQAEGRTGSGTSDADGPGGPGI.PDGPAGNAGPAGAGATGARG

PRGAGAAASGPGGAGPAGPAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

ELARRSLADAPPLPVPAGVLLKEFTVSGNLTITRLTADHRQLSLISLCLQLSLILM

MTQCELVFVLAAPSPGRR"

BASE COUNT 131 a 229 c 256 g 139 t

ORIGIN

Query Match 99.6%; Score 749; DB 9; Length 755;

Best Local Similarity 100.0%; Pred. No. 5.5e-125;

Matches 749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CTGCTGCGCCCTGACCTTCTCTTGAGAGCCGGGCAAGAGTCCGAGCCATGAGAGCCG 63

Db 1 CTGCTGCGCCCTGACCTTCTCTTGAGAGCCGGGCAAGAGTCCGAGCCATGAGAGCCG 60

QY 64 AAGCCCGGCGCAGAGGGGTTGACGCGCGATGCTGATGAGCCAGAGGCCCTGACATTC 123

Db 61 AAGCCCGGCGCAGAGGGGTTGACGCGCGATGCTGATGAGCCAGAGGCCCTGACATTC 120

QY 124 CTGATGCGCCAGAGGGGCAATGCTGCGCGCCAGAGAGAGAGAGAGAGAGAGAGAG 183

Db 121 CTGATGCGCCAGAGGGGCAATGCTGCGCGCCAGAGAGAGAGAGAGAGAGAGAGAG 180

QY 184 GTCCCGGGGCGCAGAGGGCAGCAAGGGCTGCGGGCGGGAGAGAGAGAGAGAGAGAG 243

Db 181 GTCCCGGGGCGCAGAGGGCAGCAAGGGCTGCGGGCGGGAGAGAGAGAGAGAGAGAG 240

QY 244 CGCATGCGCGCGCGGCTTCAAGGCTGATGATGATGATGATGATGATGATGATGATGAT 303

Db 241 CGCATGCGCGCGGCTTCAAGGCTGATGATGATGATGATGATGATGATGATGATGAT 300

QY 304 AGAGCGCGCTGCTTGAAGTTCTACTCGCCATGCTTTCGCGACACCCATGAGAGCAGAGC 363

Db 301 AGAGCGCGCTGCTTGAAGTTCTACTCGCCATGCTTTCGCGACACCCATGAGAGCAGAGC 360

QY 364 TGGCCCGCAGAGCCTGAGCCAGAGATGCCCAACCGCTTCCGTCAGAGGGTGTCTTGA 423

Db 361 TGGCCCGCAGAGCCTGAGCCAGAGATGCCCAACCGCTTCCGTCAGAGGGTGTCTTGA 420

QY 424 AGAGTTCACTGTGTGCGGAGCATATACATACATACATACATACATACATACATACAT 483

Db 421 AGAGTTCACTGTGTGCGGAGCATATACATACATACATACATACATACATACATACAT 480

QY 484 AACTGACGCTTCCTCATACAGCTCTCTGTCTCAGAGCTTCCCTGTGATGTGATCAGC 543

Db 481 AACTGACGCTTCCTCATACAGCTCTCTGTCTCAGAGCTTCCCTGTGATGTGATCAGC 540

QY 544 AGTGTCTTTCGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 603

Db 541 AGTGTCTTTCGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600

QY 604 CTGCGCCGCTTCTCTAGTCTAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 663

Db 601 CTGCGCCGCTTCTCTAGTCTAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660

QY 664 AGTTCATTTGGGGGCTGATTTGTTGCTGCTGAGAGAGAGGCTTATGTTTGTTC 723

Db 661 AGTTCATTTGGGGGCTGATTTGTTGCTGCTGAGAGAGAGGCTTATGTTTGTTC 720

QY 724 TGTAGAAATTAATACTGAGCTACGAAAAA 752

Db 721 TGTAGAAATTAATACTGAGCTACGAAAAA 749

Search completed: February 13, 2004, 12:39:07

Job time : 3045 secs

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 806)
AUTHORS Wang, R.-F., Johnston, S.L., Zeng, G., Topalian, S.L., Schwartzentruber, D.J., and Rosenberg, S.A.
TITLE A Breast and Melanoma-Shared Tumor Antigen: T Cell Responses to Antigenic Peptides Translated from Different Open Reading Frames
JOURNAL J. Immunol. 161, 3596-3606 (1998)
REFERENCE 2 (bases 1 to 806)
AUTHORS Wang, R.-F., and Rosenberg, S.A.
TITLE Direct Submision
JOURNAL Submitted (14-DEC-1997) Surgery Branch, National Cancer Institute, 9000 Rockville Pike, Bethesda, MD 20892, USA
FEATURES
source
1..806
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
89..631
/note="CAG-3; tumor antigen recognized by cytotoxic T lymphocytes"
/codon_start=1
/product="cancer antigen-3"
/protein_id="A005202.1"
/db_xref="GI:4104677"
/translation="MQAEGRGTSSTADADPGAPGIPDPGPNAGAPGEPAGATGGRG PRGAGARASGPGGARGPRGHGAASLNGCCCGARGAPSRILEFYIAMPATPMEA ELARSLAODAPLPVPVGVLLKEFTVSGNLTIRLTADHROQLSISGLQQLSLIM WITOCPLVPLAQPSSGQR"
123..305
/note="CAG-3-ORF2"
/codon_start=1
/product="cancer antigen-3-ORF2"
/protein_id="A005203.1"
/db_xref="GI:4159898"
/translation="MTMAQELALFLMAQMALAQAERVPRAAEVPAQGGQPPRGRE EAPRGVEMARLQG"
BASE COUNT 155 a 236 c 270 g 145 t
ORIGIN
Query Match 99.7%; Score 750; DB 9; Length 806;
Best Local Similarity 100.0%; Pred. No. 3,6e-125;
Matches 750; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

398 CTGGCCCGCAGAGCGTGGCCGAGATGCCACCGCTTCCCGTCAGGAGGCTGCTTCTG 457
QY 423 AAGAGTTCACTGTGTCCGCGACAATACATCTACTGACTGACTGAGACCAACCGC 482
DB 458 AAGAGTTCACTGTGTCCGCGACAATACATCTACTGACTGACTGAGACCAACCGC 517
QY 483 CAATGCAAGCTCTCCATCACTCTCTGTCTCCAGAGCTTTCCCTGTTGATGTGATACG 542
DB 518 CAATGCAAGCTCTCCATCACTCTCTGTCTCCAGAGCTTTCCCTGTTGATGTGATACG 577
QY 543 CAGTCTTTGCCCCGCTTTTGGCTCAGAGCTCCCTCAGGAGCAGAGCGCTAAGCCAG 602
DB 578 CAGTCTTTGCCCCGCTTTTGGCTCAGAGCTCCCTCAGGAGCAGAGCGCTAAGCCAG 637
QY 603 CTGGAGCCCTCTCTCTAGATCATGTCTCTCTCTAGGGAATGCTCCAGACAGATGCGC 662
DB 638 CTGGAGCCCTCTCTCTAGATCATGTCTCTCTCTAGGGAATGCTCCAGACAGATGCGC 697
QY 663 CAGTCTTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 722
DB 698 CAGTCTTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 757
QY 723 CTGTAGAAATTAATCTAGCTACGAAAAA 752
DB 758 CTGTAGAAATTAATCTAGCTACGAAAAA 787
RESULT 14
AR042537 755 bp DNA linear PAT 29-SEP-1999
LOCUS AR042537
DEFINITION Sequence 8 from patent US 5811519.
ACCESSION AR042537
VERSION AR042537.1 GI:5963033
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 755)
AUTHORS Lethe, B., Lucas, S., De Smet, C., Godelaine, D. and Boon-Falleur, T.
TITLE IL-1 tumor specific genes
JOURNAL Patent: US 5811519-A 8 22-SEP-1998;
FEATURES
source
1..755
/organism="unknown"
BASE COUNT 131 a 229 c 256 g 139 t
ORIGIN
Query Match 99.6%; Score 749; DB 6; Length 755;
Best Local Similarity 100.0%; Pred. No. 5.5e-125;
Matches 749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 241 GTCCGATGCGCGCGGCTTCAAGGCTGAATGATGCTGCAGATGCGGGGCCAAGGGGCG 300
Db 241 GTCCGATGCGCGCGGCTTCAAGGCTGAATGATGCTGCAGATGCGGGGCCAAGGGGCG 300
QY 301 CGAGAGCGCGCTGCTTGAATGCTTACCGCGGCGATGCTTGGCGAGACCAATGAGACAG 360
Db 301 CGAGAGCGCGCTGCTTGAATGCTTACCGCGGCGATGCTTGGCGAGACCAATGAGACAG 360
QY 361 AGCTGCGCGCGAGAGCTTGCAGAGATGCGGCGGCGGCTTCCGTCGCGAGGCGTCTTC 420
Db 361 AGCTGCGCGCGAGAGCTTGCAGAGATGCGGCGGCGGCTTCCGTCGCGAGGCGTCTTC 420
QY 421 TGAAGAGATGCTGCTGTCGCGGCGAATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 TGAAGAGATGCTGCTGTCGCGGCGAATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 481 GCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 481 GCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 541 CGCATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 541 CGCATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 601 AGCTGCGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 601 AGCTGCGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 661 GCGAGTTCATGTCGCGGCGCTGATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 661 GCGAGTTCATGTCGCGGCGCTGATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 721 TTCTGTAGAAAATAAACTGAGCTACGAAAAA 752
Db 721 TTCTGTAGAAAATAAACTGAGCTACGAAAAA 752

RESULT 12
LOCUS HS087459 752 bp mRNA linear PRI 22-DEC-1999
DEFINITION Human autoimmunogenic cancer/testis antigen NY-ESO-1 mRNA, complete cds
ACCESSION U87459.1 GI:1890098
VERSION U87459.1
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 752)
Chen, Y.-T., Scanlan, M.J., Sahin, U., Threzi, O., Guze, A.O., Teang, S.,
Williamson, B., Stockert, E., Pfundschuh, W. and Old, L.J.
A testicular antigen aberrantly expressed in human cancers detected
by autoologous antibody screening
Proc. Natl. Acad. Sci. U.S.A. 94 (5), 1914-1918 (1997)
JOURNAL
MEDLINE 97203161
PUBMED 9050879
REFERENCE 2 (bases 1 to 752)
AUTHORS Chen, Y.-T.
TITLE Direct Submission
JOURNAL Submitted (28-JAN-1997) Ludwig Institute for Cancer Research, New
York Branch, 1275 York Avenue, New York, NY 10021, USA
FEATURES
source
1..752
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="vesophageal squamous cell carcinoma expression
cDNA library"
54..596
/codon_start=1
/product="autoimmunogenic cancer/testis antigen NY-ESO-1"

```

```

BASE COUNT 126 a 230 c 256 g 140 t
ORIGIN
Query Match 100.0%; Score 752; DB 9; Length 752;
Best Local Similarity 100.0%; Pred. No. 1,6e-125;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/protein_id="AAB49693.1"
/db_xref="GI:1890098"
/translation="MOARGTGSGTDADPGPGRIIDPGGNAGSGENGATGRC
PRGAAARASGPGGAPRGPGGASGINSGCCRGARSPERLLEFTAMFATPMEA
ELARSLADADAPLPVPGVLLKEFTVSNILITRLTADHRQLQLSISSCLQSLLM
WITOCFLVFLAOPSPGQR"
QY 1 ATCCGTGCGCGCTGACCTTCTCTGAGAGCGCGGAGAGGCTCCGAGCATGACAG 60
Db 1 ATCCGTGCGCGCTGACCTTCTCTGAGAGCGCGGAGAGGCTCCGAGCATGACAG 60
QY 61 CCGAAGCGCGGCGGCAAGGCGGCTTGAAGCGGCGATGCTGATGCGCCAGAGCGCTTGGCA 120
Db 61 CCGAAGCGCGGCGGCAAGGCGGCTTGAAGCGGCGATGCTGATGCGCCAGAGCGCTTGGCA 120
QY 121 TTCTGATGCGCCAGGCGGCAATGCTGCGCGGCGCCAGAGAGCGGCTGCAAGCGGCA 180
Db 121 TTCTGATGCGCCAGGCGGCAATGCTGCGCGGCGCCAGAGAGCGGCTGCAAGCGGCA 180
QY 181 GAGGTCCCGGCGGCGGAGGCGGAGGAGGCTGCGGCGCGGAGAGGCGCGCGGCG 240
Db 181 GAGGTCCCGGCGGCGGAGGCGGAGGAGGAGGCTGCGGCGCGGAGAGGCGCGCGGCG 240
QY 241 GTCCGATGCGCGCGCGGCTTCAAGGCTGAAATGATGCTGCGAGTGCAGGCGGCGGCG 300
Db 241 GTCCGATGCGCGCGCGGCTTCAAGGCTGAAATGATGCTGCGAGTGCAGGCGGCGGCG 300
QY 301 CGAGAGCGCGCTGCTTGAATGCTTCACTGCGCATGCTTTCGCGACACCATGAAAGAG 360
Db 301 CGAGAGCGCGCTGCTTGAATGCTTCACTGCGCATGCTTTCGCGACACCATGAAAGAG 360
QY 361 AGCTGCGCGGAGAGCTTGGCGGAGAGGCGGCGGCTTCCGTCGCGAGGCGTTC 420
Db 361 AGCTGCGCGGAGAGCTTGGCGGAGAGGCGGCGGCTTCCGTCGCGAGGCGTTC 420
QY 421 TGAAGAGATGCTGCTGCTGCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 TGAAGAGATGCTGCTGCTGCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 481 GCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 481 GCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 541 CGAGAGCTTCTGCGCGGCTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 541 CGAGAGCTTCTGCGCGGCTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 601 AGCTGCGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 601 AGCTGCGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 661 GCGAGTTCATGTCGCGGCGCTGATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 661 GCGAGTTCATGTCGCGGCGCTGATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 721 TTCTGTAGAAAATAAACTGAGCTACGAAAAA 752
Db 721 TTCTGTAGAAAATAAACTGAGCTACGAAAAA 752

RESULT 13
LOCUS AF038567 806 bp mRNA linear PRI 16-JAN-1999
DEFINITION Homo sapiens cancer antigen-3 and cancer antigen-3-ORF2 mRNA,
complete cds.
ACCESSION AF038567
VERSION AF038567.1 GI:4104676

```



```

AX568285      752 bp      DNA      linear      PAT 22-MAR-2003
LOCUS         AX568285
DEFINITION    Sequence 201 from Patent WO03000928.
ACCESSION     AX568285
VERSION       AX568285.1 GI:29160798
KEYWORDS
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
  1. Poulsen, H.S., Pedersen, N., Mortensen, S., Sorensen, S.B.,
    Petersen, M.W. and Elsen, H.I.
    Methods for identification of cancer cell surface molecules and
    cancer specific promoters, and therapeutic uses thereof
    Patent: WO 03000928-A 201 03-JAN-2003;
    Odin Medical A/S (DK)

FEATURES
  source
    1..752
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"

BASE COUNT    126 a      230 c      256 g      140 t

ORIGIN
Query Match      100.0%; Score 752; DB 6; Length 752;
Best Local Similarity 100.0%; Pred. No. 1,6e-125;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTCGTGGGCGCTTGAACCTTCTCTGAGAGCGGCGAGAGGCTTCGAGGCATGCAAG 60
DB 1 ATCTCGTGGGCGCTTGAACCTTCTCTGAGAGCGGCGAGAGGCTTCGAGGCATGCAAG 60
QY 61 CGAAGGCGGCGGCAAGGAGGTTGACGCGGCGATGCTGATGCGCCAGAGGCGCTGGCA 120
DB 61 CGAAGGCGGCGGCAAGGAGGTTGACGCGGCGATGCTGATGCGCCAGAGGCGCTGGCA 120
QY 121 TTCTGATGAGCCGAGGCGCAATGCTGCGCGCCAGAGAGGCGGGTGCACAGGCGCA 180
DB 121 TTCTGATGAGCCGAGGCGCAATGCTGCGCGCCAGAGAGGCGGGTGCACAGGCGCA 180
QY 121 TTCTGATGAGCCGAGGCGCAATGCTGCGCGCCAGAGAGGCGGGTGCACAGGCGCA 180
DB 121 TTCTGATGAGCCGAGGCGCAATGCTGCGCGCCAGAGAGGCGGGTGCACAGGCGCA 180
QY 181 GAGGTCCCGGCGGCGAGGCGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
DB 181 GAGGTCCCGGCGGCGAGGCGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
QY 241 GTCCGATGAGCGGCGGCGGCTTCAGGCGCTGATGAGATGCGGCGGCGAGGCGGCGG 300
DB 241 GTCCGATGAGCGGCGGCGGCTTCAGGCGCTGATGAGATGCGGCGGCGAGGCGGCGG 300
QY 301 CGAAGGCGGCGGCGGCTTCAGGCGCTTCAGGCGCTTCAGGCGCTTCAGGCGCTTCAG 360
DB 301 CGAAGGCGGCGGCGGCTTCAGGCGCTTCAGGCGCTTCAGGCGCTTCAGGCGCTTCAG 360
QY 361 AGCTGCGCGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
DB 361 AGCTGCGCGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
QY 421 TGAAGAGTTCACTGTGTCCGCACTAATGCTATCCGACTGACTGCTGCAAGCCACC 480
DB 421 TGAAGAGTTCACTGTGTCCGCACTAATGCTATCCGACTGACTGCTGCAAGCCACC 480
QY 481 GCGAATGAGGCTCTCATAGCTCTGTCTCAGAGAGGCTTCGCGGTTGATGATGATGAT 540
DB 481 GCGAATGAGGCTCTCATAGCTCTGTCTCAGAGAGGCTTCGCGGTTGATGATGATGAT 540
QY 541 CGAGTGTCTTCTGCGCGGTTTGTGCTGACGCTTCCTTCAAGGCGAGGCGCTAAGCCC 600
DB 541 CGAGTGTCTTCTGCGCGGTTTGTGCTGACGCTTCCTTCAAGGCGAGGCGCTAAGCCC 600
QY 601 AGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
DB 601 AGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660

```

```

QY 661 GCGAGTTCATTGTGGGGGCGCTGATTTGTTGTGCTGAGAGAGGAGCGCTTACATGTTGT 720
DB 661 GCGAGTTCATTGTGGGGGCGCTGATTTGTTGTGCTGAGAGAGGAGCGCTTACATGTTGT 720
QY 721 TTCTGTAGAAAATPAACTGAGCTACGAAAAA 752
DB 721 TTCTGTAGAAAATPAACTGAGCTACGAAAAA 752

RESULT 11
LOCUS     BD069515
DEFINITION Isolated nucleic acid molecule encoding cancer associated antigen,
            the antigen itself, and uses thereof.
ACCESSION BD069515
VERSION    BD069515.1 GI:22615118
KEYWORDS   JP 2001507212-A/1.
SOURCE     unidentified
ORGANISM   unidentified
            unclassified.

REFERENCE
  1. (bases 1 to 752)
  Chen, F.T., Scanlan, M., Gure, A., Old, L.J., Jager, E., Knuth, A. and
  Drifflout, J.W.
  Isolated nucleic acid molecule encoding cancer associated antigen,
  the antigen itself, and uses thereof
  Patent: JP 2001507212-A 1 05-JUN-2001;
  LUDWIG INSTITUTE FOR CANCER RESEARCH

COMMENT
  OS Unidentified
  PN JP 2001507212-A/1
  PD 05-JUN-2001
  PF 15-SEP-1997 JP 1998516567
  PR 03-OCT-1996 US 08/725182
  PI YAO TSENG CHEN, MATTHEW SCANLAN, ALI GURE, LLOYD J OLD, ELKE
  JAGER.
  PT ALEXANDER KNUTH, TAN W DRIFFOUT
  PC C07H21/04, C07K14/435, C07K16/18, C12N15/64, C12N15/85, C12Q1/68,
  PC C12P19/34,
  PC G01N33/53, A61K35/00, A61K38/43, A61K48/00
  CC Strandedness: Double;
  CC Topology: Linear;
  CC Isolated nucleic acid molecule encoding cancer associated CC
  CC antigen, the
  CC antigen itself, and uses thereof
  FH Key
  FT source
  FT Location/Qualifiers
  1..752
  /organism="unidentified"
  /mol_type="genomic DNA"
  /db_xref="taxon:32644"

BASE COUNT    126 a      230 c      256 g      140 t

ORIGIN
Query Match      100.0%; Score 752; DB 6; Length 752;
Best Local Similarity 100.0%; Pred. No. 1,6e-125;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTCGTGGGCGCTTGAACCTTCTCTGAGAGCGGCGAGAGGCTTCGAGGCATGCAAG 60
DB 1 ATCTCGTGGGCGCTTGAACCTTCTCTGAGAGCGGCGAGAGGCTTCGAGGCATGCAAG 60
QY 61 CGAAGGCGGCGGCAAGGAGGTTGACGCGGCGATGCTGATGCGCCAGAGGCGCTGGCA 120
DB 61 CGAAGGCGGCGGCAAGGAGGTTGACGCGGCGATGCTGATGCGCCAGAGGCGCTGGCA 120
QY 121 TTCTGATGAGCCGAGGCGCAATGCTGCGCGCCAGAGAGGCGGGTGCACAGGCGCA 180
DB 121 TTCTGATGAGCCGAGGCGCAATGCTGCGCGCCAGAGAGGCGGGTGCACAGGCGCA 180
QY 121 TTCTGATGAGCCGAGGCGCAATGCTGCGCGCCAGAGAGGCGGGTGCACAGGCGCA 180
DB 121 TTCTGATGAGCCGAGGCGCAATGCTGCGCGCCAGAGAGGCGGGTGCACAGGCGCA 180
QY 181 GAGGTCCCGGCGGCGAGGCGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
DB 181 GAGGTCCCGGCGGCGAGGCGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 126 a 230 c 256 g 140 t
ORIGIN

Query Match 100.0%; Score 752; DB 6; Length 752;
Best Local Similarity 100.0%; Pred. No. 1.6e-125; Indels 0; Gaps 0;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTCTGAGGCGCCCTGACCTTCTCTCTGAGAGCCGCGGCAAGAGCTCCGAGCCATGCAAG 60
DB 1 ATCTCTGAGGCGCCCTGACCTTCTCTCTGAGAGCCGCGGCAAGAGCTCCGAGCCATGCAAG 60
QY 61 CCGAAGCGCGGCGCAAGAGGCGTTGAGAGGCGGATGCTGATGCGCCAGAGGCGCTTGGCA 120
DB 61 CCGAAGCGCGGCGCAAGAGGCGTTGAGAGGCGGATGCTGATGCGCCAGAGGCGCTTGGCA 120
QY 121 TTCTGATGCGCCAGAGGCGCAATGCTGCGGCGCCAGAGAGCGGATGCCAAGCGGCGCA 180
DB 121 TTCTGATGCGCCAGAGGCGCAATGCTGCGGCGCCAGAGAGCGGATGCCAAGCGGCGCA 180
QY 181 GAGGTCCCGGCGCGCAAGAGGCGCTGAGAGGCGCTGAGAGGCGCGCGCGCG 240
DB 181 GAGGTCCCGGCGCGCAAGAGGCGCTGAGAGGCGCTGAGAGGCGCGCGCGCG 240
QY 241 GTCCGATGCGCGCGCGCTTCAAGAGGCTGAATGATGCTGAGATGCGGCGCGCAAGGCGCG 300
DB 241 GTCCGATGCGCGCGCGCTTCAAGAGGCTGAATGATGCTGAGATGCGGCGCGCAAGGCGCG 300
QY 301 CGAGAGCGCGCGCTTCAAGATGCTGCGGCGCTTTCGCGACACCCATGGAAGCA 360
DB 301 CGAGAGCGCGCGCTTCAAGATGCTGCGGCGCTTTCGCGACACCCATGGAAGCA 360
QY 361 AGCTGCGCGCGCGCTTCAAGATGCTGCGGCGCTTTCGCGACACCCATGGAAGCA 420
DB 361 AGCTGCGCGCGCGCTTCAAGATGCTGCGGCGCTTTCGCGACACCCATGGAAGCA 420
QY 421 TGAAGAGTTCATCTGTGTCGCGCAACATGCTGATGCTGCGGCGCTTTCGCGACACCC 480
DB 421 TGAAGAGTTCATCTGTGTCGCGCAACATGCTGATGCTGCGGCGCTTTCGCGACACCC 480
QY 481 GCGAAGTGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 481 GCGAAGTGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 541 GCGAAGTGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 541 GCGAAGTGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 601 AGCTGCGCGCGCGCTTCAAGATGCTGCGGCGCTTTCGCGACACCCATGGAAGCA 660
DB 601 AGCTGCGCGCGCGCTTCAAGATGCTGCGGCGCTTTCGCGACACCCATGGAAGCA 660
QY 661 GCGAAGTGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 GCGAAGTGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 TTCTGTAAGAAATTAATACTGAGCTACGAAAAA 752
DB 721 TTCTGTAAGAAATTAATACTGAGCTACGAAAAA 752

RESULT 9
AX235020 752 bp DNA linear PAT 11-SEP-2001
LOCUS AX235020
DEFINITION Sequence 1 from Patent WO0162917.
ACCESSION AX235020
VERSION AX235020.1 GI:15593677
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
1 Lettre, B. and Boon-Falleur, T.
AUTHORS Isolated genomic sequences which encode the ny-eso-1 cancer/testis
TITLE tumor antigen and uses thereof
JOURNAL Patent: WO 0162917-A 1 30-AUG-2001;
LUDWIG INSTITUTE FOR CANCER RESEARCH (US)
FEATURES
source location/Qualifiers
1..752
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 126 a 230 c 256 g 140 t
ORIGIN

Query Match 100.0%; Score 752; DB 6; Length 752;
Best Local Similarity 100.0%; Pred. No. 1.6e-125; Indels 0; Gaps 0;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTCTGAGGCGCCCTGACCTTCTCTCTGAGAGCCGCGGCAAGAGCTCCGAGCCATGCAAG 60
DB 1 ATCTCTGAGGCGCCCTGACCTTCTCTCTGAGAGCCGCGGCAAGAGCTCCGAGCCATGCAAG 60
QY 61 CCGAAGCGCGGCGCAAGAGGCGTTGAGAGGCGGATGCTGATGCGCCAGAGGCGCTTGGCA 120
DB 61 CCGAAGCGCGGCGCAAGAGGCGTTGAGAGGCGGATGCTGATGCGCCAGAGGCGCTTGGCA 120
QY 121 TTCTGATGCGCCAGAGGCGCAATGCTGCGGCGCCAGAGAGCGGATGCCAAGCGGCGCA 180
DB 121 TTCTGATGCGCCAGAGGCGCAATGCTGCGGCGCCAGAGAGCGGATGCCAAGCGGCGCA 180
QY 181 GAGGTCCCGGCGCGCAAGAGGCGCTGAGAGGCGCTGAGAGGCGCGCGCGCG 240
DB 181 GAGGTCCCGGCGCGCAAGAGGCGCTGAGAGGCGCTGAGAGGCGCGCGCGCG 240
QY 241 GTCCGATGCGCGCGCGCTTCAAGAGGCTGAATGATGCTGAGATGCGGCGCGCAAGGCGCG 300
DB 241 GTCCGATGCGCGCGCGCTTCAAGAGGCTGAATGATGCTGAGATGCGGCGCGCAAGGCGCG 300
QY 301 CGAGAGCGCGCGCTTCAAGATGCTGCGGCGCTTTCGCGACACCCATGGAAGCA 360
DB 301 CGAGAGCGCGCGCTTCAAGATGCTGCGGCGCTTTCGCGACACCCATGGAAGCA 360
QY 361 AGCTGCGCGCGCGCTTCAAGATGCTGCGGCGCTTTCGCGACACCCATGGAAGCA 420
DB 361 AGCTGCGCGCGCGCTTCAAGATGCTGCGGCGCTTTCGCGACACCCATGGAAGCA 420
QY 421 TGAAGAGTTCATCTGTGTCGCGCAACATGCTGATGCTGCGGCGCTTTCGCGACACCC 480
DB 421 TGAAGAGTTCATCTGTGTCGCGCAACATGCTGATGCTGCGGCGCTTTCGCGACACCC 480
QY 481 GCGAAGTGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 481 GCGAAGTGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 541 GCGAAGTGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 541 GCGAAGTGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 601 AGCTGCGCGCGCGCTTCAAGATGCTGCGGCGCTTTCGCGACACCCATGGAAGCA 660
DB 601 AGCTGCGCGCGCGCTTCAAGATGCTGCGGCGCTTTCGCGACACCCATGGAAGCA 660
QY 661 GCGAAGTGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 GCGAAGTGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 TTCTGTAAGAAATTAATACTGAGCTACGAAAAA 752
DB 721 TTCTGTAAGAAATTAATACTGAGCTACGAAAAA 752

RESULT 10

```

Db      121 TTCTCATGCGCCAGAGGCGCATGTGCGGCCAGAGAGGCGGCTGCCAGCGGCGGCA 180
QY      181 GAGGTCCCCCGGCGCGCAGAGGCGCATGAGGCGCTCGGAGAGAGCGCGCGCGG 240
Db      181 GAGGTCCCCCGGCGCGCAGAGGCGCATGAGGCGCTCGGAGAGAGCGCGCGG 240
QY      241 GTCCGATGCGCGCGCGCTTCAAGGCTGATGATGCTGAGATGCGGCGCGCGG 300
Db      241 GTCCGATGCGCGCGCGCTTCAAGGCTGATGATGCTGAGATGCGGCGCGG 300
QY      301 CGAGAGCGCGCGCTTGAATTTCACTGCGCATGCTTTCGCGACACCCATGGAAGCAG 360
Db      301 CGAGAGCGCGCGCTTGAATTTCACTGCGCATGCTTTCGCGACACCCATGGAAGCAG 360
QY      361 AGCTGGCGCGCAGAGGCGCTGAGAGATGCGCGCATGCTTTCGCGACAGGCGGCTTC 420
Db      361 AGCTGGCGCGCAGAGGCGCTGAGAGATGCGCGCATGCTTTCGCGACAGGCGGCTTC 420
QY      421 TGAAGAGTTCACTGTGTCCGCAACATCTGACTATCCGACTGACTGCTGCAAGCACC 480
Db      421 TGAAGAGTTCACTGTGTCCGCAACATCTGACTATCCGACTGACTGCTGCAAGCACC 480
QY      481 GCCAATGCAAGCTCTCCATGAGCTCCGCTGCGAGAGCTTTCGCTTGAATGAGATCA 540
Db      481 GCCAATGCAAGCTCTCCATGAGCTCCGCTGCGAGAGCTTTCGCTTGAATGAGATCA 540
QY      541 CGCAGTCTTCTGCGCGGTGTTTGGCTCAGCTCCCTCGAGGCGAGCGCGTAAGCCC 600
Db      541 CGCAGTCTTCTGCGCGGTGTTTGGCTCAGCTCCCTCGAGGCGAGCGCGTAAGCCC 600
QY      601 AGCTGGCGCGCTTCTGAGTCAATGCTCTCCCTGAGGAGATGCTCCAGAGAGAGG 660
Db      601 AGCTGGCGCGCTTCTGAGTCAATGCTCTCCCTGAGGAGATGCTCCAGAGAGAGG 660
QY      661 GCCAGTTCAATGTGGGGGCTGATTTGTTGCTGCTGAGAGAGAGCGGCTTACATGTTGT 720
Db      661 GCCAGTTCAATGTGGGGGCTGATTTGTTGCTGCTGAGAGAGAGCGGCTTACATGTTGT 720
QY      721 TTCTGTAGAAAATAAACTGAGCTACGAAAAA 752
Db      721 TTCTGTAGAAAATAAACTGAGCTACGAAAAA 752

```

RESULT 7

```

AX079857 752 bp DNA linear PAT 22-FEB-2001
LOCUS      AX079857
DEFINITION Sequence 1 from Patent WO0107917.
ACCESSION  AX079857
VERSION     AX079857.1 GI:13159392
KEYWORDS
SOURCE
ORGANISM   Homo sapiens (human)
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

REFERENCE

```

AUTHORS    Jeger,E., Stockert,E., Old,L.J., Knuth,A., Chen,Y.T. and Scanlan,M.
TITLE      Determining antibodies to my-eso-1 in cancer patients
JOURNAL    Patent: WO 0107917-A 1 01-FEB-2001;
            LUDWIG INSTITUTE FOR CANCER RESEARCH (US) ; MEMORIAL
            SLOAN-KETTERING CANCER CENTER (US) ; Cornell Research Foundation
            (US)

```

FEATURES

```

source      Location/Qualifiers
            1..752
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
BASE COUNT  126 a 230 c 256 g 140 t
ORIGIN

```

```

Query Match      100.0%; Score 752; DB 6; Length 752;
Best Local Similarity 100.0%; Pred. No. 1,6e-125;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 ATCTCTGTGGCGCTTGAATCTTCTGTAGAGCCGCGCAGAGGCTCCGAGCCATGCAAG 60
Db      1 ATCTCTGTGGCGCTTGAATCTTCTGTAGAGCCGCGCAGAGGCTCCGAGCCATGCAAG 60
QY      61 CCGAAGCGCGCGCGCAGAGGCGCTTCAAGGCGCATGCTGATGAGCCAGAGGCGCTTGGCA 120
Db      61 CCGAAGCGCGCGCGCAGAGGCGCTTCAAGGCGCATGCTGATGAGCCAGAGGCGCTTGGCA 120
QY      121 TTCTGTATGCGCCAGAGGCGCGCATGCTGCGCGCCAGAGAGAGCGCGGATGCCACGCGGCA 180
Db      121 TTCTGTATGCGCCAGAGGCGCGCATGCTGCGCGCCAGAGAGAGCGCGGATGCCACGCGGCA 180
QY      181 GAGGTCCCCCGGCGCGCAGAGGCGCATGAGGCGCTTGGGGGCGCGGAGAGAGGCGCGG 240
Db      181 GAGGTCCCCCGGCGCGCAGAGGCGCATGAGGCGCTTGGGGGCGCGGAGAGAGGCGCGG 240
QY      241 GTCCGATGCGCGCGCGCTTCAAGGCTGATGAGATGCTGAGATGCGGCGCGCAGAGGCGG 300
Db      241 GTCCGATGCGCGCGCGCTTCAAGGCTGATGAGATGCTGAGATGCGGCGCGCAGAGGCGG 300
QY      301 CGAGAGCGCGCGCTTGAATTTCACTGCGCATGCTTTCGCGACACCCATGGAAGCAG 360
Db      301 CGAGAGCGCGCGCTTGAATTTCACTGCGCATGCTTTCGCGACACCCATGGAAGCAG 360
QY      361 AGCTGGCGCGCAGAGGCGCTGAGAGATGCGCGCATGCTTTCGCGACAGGCGGCTTC 420
Db      361 AGCTGGCGCGCAGAGGCGCTGAGAGATGCGCGCATGCTTTCGCGACAGGCGGCTTC 420
QY      421 TGAAGAGTTCACTGTGTCCGCAACATCTGACTATCCGACTGACTGCTGCAAGCACC 480
Db      421 TGAAGAGTTCACTGTGTCCGCAACATCTGACTATCCGACTGACTGCTGCAAGCACC 480
QY      481 GCCAATGCAAGCTCTCCATGAGCTCCGCTGCGAGAGCTTTCGCTTGAATGAGATCA 540
Db      481 GCCAATGCAAGCTCTCCATGAGCTCCGCTGCGAGAGCTTTCGCTTGAATGAGATCA 540
QY      541 CGCAGTCTTCTGCGCGGTGTTTGGCTCAGCTCCCTCGAGGCGAGCGCGTAAGCCC 600
Db      541 CGCAGTCTTCTGCGCGGTGTTTGGCTCAGCTCCCTCGAGGCGAGCGCGTAAGCCC 600
QY      601 AGCTGGCGCGCTTCTGAGTCAATGCTCTCCCTGAGGAGATGCTCCAGAGAGAGG 660
Db      601 AGCTGGCGCGCTTCTGAGTCAATGCTCTCCCTGAGGAGATGCTCCAGAGAGAGG 660
QY      661 GCCAGTTCAATGTGGGGGCTGATTTGTTGCTGCTGAGAGAGAGCGGCTTACATGTTGT 720
Db      661 GCCAGTTCAATGTGGGGGCTGATTTGTTGCTGCTGAGAGAGAGCGGCTTACATGTTGT 720
QY      721 TTCTGTAGAAAATAAACTGAGCTACGAAAAA 752
Db      721 TTCTGTAGAAAATAAACTGAGCTACGAAAAA 752

```

RESULT 8

```

AX107123 752 bp DNA linear PAT 30-APR-2001
LOCUS      AX107123
DEFINITION Sequence 1 from Patent WO0123560.
ACCESSION  AX107123
VERSION     AX107123.1 GI:13922629
KEYWORDS
SOURCE
ORGANISM   Homo sapiens (human)
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

REFERENCE

```

AUTHORS    Tureci,O., Sahin,U. and Pfreundschuh,M.
TITLE      Isolated peptides which bind to mhc class ii molecules, and uses
JOURNAL    Patent: WO 0123560-A 1 05-APR-2001;
            LUDWIG INSTITUTE FOR CANCER RESEARCH (US)

```

FEATURES

```

source      Location/Qualifiers
            1..752

```

QY 721 TTCTGTGAGAAATATAAAGTGAAGCTACGAAAA 752
 Db 721 TTCTGTGAGAAATATAAAGTGAAGCTACGAAAA 752

RESULT 5
 AX024691 752 bp DNA linear PAT 15-SEP-2000
 LOCUS Sequence 7 from Patent EP1001022.
 DEFINITION AX024691
 ACCSSION AX024691
 VERSION AX024691.1 GI:10184777
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Aarnoudse, C.A. and Schrier, P.
 Camel, an alternative translation product of the tumour antigen
 1age-1
 Patent: EP 1001022-A 7 17-MAY-2000;
 BOHRINGER INGELHEIM INT (DE) ; UNIV HOSPITAL LEIDEN (NL)

JOURNAL
 Location/Qualifiers
 1. 752
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 1. 53
 54. 596
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC08999.1"
 /db_xref="GI:10184778"
 /translation="MOAEGRGSTGSDADPGPGGIPDPGGNAGGPGHAGATGGRG
 PRGGAARASGPGGAGPRGPGGAGAGNGCCRCGAGPESRLLEFLAMPFATPMEA
 ELARRSLAODAPPLPVGVLKEFTVSGNLTITLTADHRQLSTSSCLOQLSLIM
 WITOCFLPVPFLAQPSSGGR"

5'UTR
 CDS
 3'UTR
 BASE COUNT 126 a 230 c 256 g 140 t
 ORIGIN

Query Match 100.0%; Score 752; DB 6; Length 752;
 Best Local Similarity 100.0%; Pred. No. 1.6e-125; Indels 0; Gaps 0;
 Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTGTGAGGCGCTTCTCTCTGAGAGCGCGGCAAGGCTCCGAGACCATGCAAG 60
 Db 1 ATCTGTGAGGCGCTTCTCTCTGAGAGCGCGGCAAGGCTCCGAGACCATGCAAG 60

QY 61 CCGAAGGCGCGGCGCAAGGCGGCTTGAAGGCGGATGATGAGCGCCAGAGGCGCTGCA 120
 Db 61 CCGAAGGCGCGGCGCAAGGCGGCTTGAAGGCGGATGATGAGCGCCAGAGGCGCTGCA 120

QY 121 TTCTGTATGCGCCAGGCGGCAATGCTGCGCGGCCAGAGAGAGCGGCTGCCACGCGCGCA 180
 Db 121 TTCTGTATGCGCCAGGCGGCAATGCTGCGCGGCCAGAGAGAGCGGCTGCCACGCGCGCA 180

QY 181 GAGGTCCCGCGGCGCGCAAGGCGGCTTGAAGGCGGCTTGAAGGCGGCGCGCGCGG 240
 Db 181 GAGGTCCCGCGGCGCGCAAGGCGGCTTGAAGGCGGCTTGAAGGCGGCGCGCGCGG 240

QY 241 GTCCGATGCGCGCGCGCTTGAAGGCGGCTTGAAGGCGGCGCGCGCGCGCGCGG 300
 Db 241 GTCCGATGCGCGCGCGCTTGAAGGCGGCTTGAAGGCGGCGCGCGCGCGCGCGG 300

QY 301 CGAAGAGCGCGCGCTTGAAGGCGGCTTGAAGGCGGCGCGCGCGCGCGCGCGG 360
 Db 301 CGAAGAGCGCGCGCTTGAAGGCGGCTTGAAGGCGGCGCGCGCGCGCGCGCGG 360

QY 361 AGCTGCGCGCGAGAGCGCTGCGCGCGAGATGCGCGCGCGCGCGCGCGCGCGG 420
 Db 361 AGCTGCGCGCGAGAGCGCTGCGCGCGAGATGCGCGCGCGCGCGCGCGCGG 420

QY 421 TGAAGAGTTCAGTGTGCGCGCAACATGACTATCCGATGCTGCGAGACCA 480
 Db 421 TGAAGAGTTCAGTGTGCGCGCAACATGACTATCCGATGCTGCGAGACCA 480

QY 481 GCCAATGAGCTTCCATCATGCTCTGCTCCAGAGCGCTTCCCTGTATGATGATCA 540
 Db 481 GCCAATGAGCTTCCATCATGCTCTGCTCCAGAGCGCTTCCCTGTATGATGATCA 540

QY 541 CGCAATGCTTCTGCGCGCGCTTGTGAGTCAAGCTTCCCTGAGAGCGCGCTTACGCC 600
 Db 541 CGCAATGCTTCTGCGCGCGCTTGTGAGTCAAGCTTCCCTGAGAGCGCGCTTACGCC 600

QY 601 AGCTGCGCGCGCTTCTGAGTCAAGCTTCCCTGAGAGCGCGCTTACGCC 660
 Db 601 AGCTGCGCGCGCTTCTGAGTCAAGCTTCCCTGAGAGCGCGCTTACGCC 660

QY 661 GCCAGTTCATGTTGCGGCGCTGATGTTGTGCTGAGAGAGAGCGCTTACATGTTGT 720
 Db 661 GCCAGTTCATGTTGCGGCGCTGATGTTGTGCTGAGAGAGAGCGCTTACATGTTGT 720

QY 721 TTCTGTGAGAAATATAAAGTGAAGCTACGAAAA 752
 Db 721 TTCTGTGAGAAATATAAAGTGAAGCTACGAAAA 752

RESULT 6
 AX024693 752 bp DNA linear PAT 15-SEP-2000
 LOCUS Sequence 9 from Patent EP1001022.
 DEFINITION AX024693
 ACCSSION AX024693
 VERSION AX024693.1 GI:10184779
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Aarnoudse, C.A. and Schrier, P.
 Camel, an alternative translation product of the tumour antigen
 1age-1
 Patent: EP 1001022-A 9 17-MAY-2000;
 BOHRINGER INGELHEIM INT (DE) ; UNIV HOSPITAL LEIDEN (NL)

JOURNAL
 Location/Qualifiers
 1. 752
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 94. 270
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC09000.1"
 /db_xref="GI:10184780"
 /translation="MIMAGALAFMAAGSMUAQGRVRAAEVPGAQGGGPRGRE
 EAPRGVRAARLQG"

5'UTR
 CDS
 3'UTR
 BASE COUNT 126 a 230 c 256 g 140 t
 ORIGIN

Query Match 100.0%; Score 752; DB 6; Length 752;
 Best Local Similarity 100.0%; Pred. No. 1.6e-125; Indels 0; Gaps 0;
 Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTGTGAGGCGCTTCTCTCTGAGAGCGCGGCAAGGCTCCGAGACCATGCAAG 60
 Db 1 ATCTGTGAGGCGCTTCTCTCTGAGAGCGCGGCAAGGCTCCGAGACCATGCAAG 60

QY 61 CCGAAGGCGCGGCGCAAGGCGGCTTGAAGGCGGATGATGAGCGCCAGAGGCGCTGCA 120
 Db 61 CCGAAGGCGCGGCGCAAGGCGGCTTGAAGGCGGATGATGAGCGCCAGAGGCGCTGCA 120

QY 121 TTCTGTATGCGCCAGGCGGCAATGCTGCGCGGCCAGAGAGAGCGGCTGCCACGCGCGCA 180

VERSION AR161582.1 GI:16227523
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 752)
 TITLE Unclassified.
 AUTHORS Chen, Y.-T., Scanlan, M., Gure, A. and Old, L.J.
 Isolated nucleic acid molecule encoding an esophageal cancer
 associated antigen
 JOURNAL Patent: US 625470-A 1 03-JUL-2001;
 FEATURES
 source Location/Qualifiers
 1..752
 /organism="unknown"
 BASE COUNT 126 a 230 c 256 g 140 t
 ORIGIN
 Query Match 100.0%; Score 752; DB 6; Length 752;
 Best Local Similarity 100.0%; Pred. No. 1.6e-125;
 Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATCTCGTGGGCGCTGACCTTCTCTCTGAGAGCGCGGACAGGCTCCGAGCCATGACAG 60
 DB 1 ATCTCGTGGGCGCTGACCTTCTCTCTGAGAGCGCGGACAGGCTCCGAGCCATGACAG 60
 QY 61 CCGAAGCGCGGACAGGCGGCTTGAAGCGGCGATGCTGATGCGCCAGAGGCGCTTGACA 120
 DB 61 CCGAAGCGCGGACAGGCGGCTTGAAGCGGCGATGCTGATGCGCCAGAGGCGCTTGACA 120
 QY 121 TTCTGTATGGCCCAAGGCGGCGAATGCTGGCGGCGCCAGAGAGCGGCTGCCACGCGGCGCA 180
 DB 121 TTCTGTATGGCCCAAGGCGGCGAATGCTGGCGGCGCCAGAGAGCGGCTGCCACGCGGCGCA 180
 QY 121 TTCTGTATGGCCCAAGGCGGCGAATGCTGGCGGCGCCAGAGAGCGGCTGCCACGCGGCGCA 180
 DB 121 TTCTGTATGGCCCAAGGCGGCGAATGCTGGCGGCGCCAGAGAGCGGCTGCCACGCGGCGCA 180
 QY 181 GAGGTCCCGGCGGCGCAAGGCGGCGCTCGGCGCGGAGAGAGCGCGCCCGCGG 240
 DB 181 GAGGTCCCGGCGGCGCAAGGCGGCGCTCGGCGCGGAGAGAGCGCGCCCGCGG 240
 QY 241 GTCCGATGGCGGCGGCGCTTCAAGGCTGATGATGCTGCAATGCGGCGCGGCGGCGGCG 300
 DB 241 GTCCGATGGCGGCGGCGCTTCAAGGCTGATGATGCTGCAATGCGGCGCGGCGGCGGCG 300
 QY 301 CGAGAGCGCGCTGCTGATGCTGATGCTGCTGCGGAGACCCATGGAAGAGAG 360
 DB 301 CGAGAGCGCGCTGCTGATGCTGATGCTGCTGCGGAGACCCATGGAAGAGAG 360
 QY 361 AGCTGGCGCGGAGGCTTGCAGGATGCCCAAGGCTTCCGTCGCAAGGCGGCTTTC 420
 DB 361 AGCTGGCGCGGAGGCTTGCAGGATGCCCAAGGCTTCCGTCGCAAGGCGGCTTTC 420
 QY 421 TGAAGAGGTTCACTGTGTCCGCGCAACAATGACTGATGCTGATGCTGCGAGACAC 480
 DB 421 TGAAGAGGTTCACTGTGTCCGCGCAACAATGACTGATGCTGATGCTGCGAGACAC 480
 QY 481 GCCAAGTCAAGTCTCCATGAGCTCTGCTCCAGAGAGCTTCCCTGTGATGATGATCA 540
 DB 481 GCCAAGTCAAGTCTCCATGAGCTCTGCTCCAGAGAGCTTCCCTGTGATGATGATCA 540
 QY 541 GCGAGTCTTTTTCGCGCGCTTTTTCGCTCAAGCTCTCCAGAGAGAGGCGCTAAGGCC 600
 DB 541 GCGAGTCTTTTTCGCGCGCTTTTTCGCTCAAGCTCTCCAGAGAGAGGCGCTAAGGCC 600
 QY 601 AGCTGGCGCGCTTCTTAAGTCAATGCTCTCCAGAGAGAGGCGCTAAGGCC 660
 DB 601 AGCTGGCGCGCTTCTTAAGTCAATGCTCTCCAGAGAGAGGCGCTAAGGCC 660
 QY 661 GCCAGTCTTTTTCGCGCGCTTTTTCGCTCAAGCTCTCCAGAGAGAGGCGCTAAGGCC 720
 DB 661 GCCAGTCTTTTTCGCGCGCTTTTTCGCTCAAGCTCTCCAGAGAGAGGCGCTAAGGCC 720
 QY 721 TTCTGTAGAAAATTAAGTCAAGTCAAGAAA 752
 DB 721 TTCTGTAGAAAATTAAGTCAAGTCAAGAAA 752

RESULT 4
 AR164507
 LOCUS AR164507 752 bp DNA linear PAT 17-OCT-2001
 DEFINITION Sequence 1 from patent US 6274145.
 ACCESSION AR164507
 VERSION AR164507.1 GI:16237557
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 752)
 TITLE Unclassified.
 AUTHORS Chen, Y.-T., Scanlan, M., Gure, A., Old, L.J., Jager, E., Alexander, K.,
 Drifflough, J.W., Tureci, O., Sahin, U. and Pfreundschuh, M.
 Isolated nucleic acid molecule encoding cancer associated antigen.
 the antigen itself, and uses thereof
 JOURNAL Patent: US 6274145-A 1 14-AUG-2001;
 FEATURES
 source Location/Qualifiers
 1..752
 /organism="unknown"
 BASE COUNT 126 a 230 c 256 g 140 t
 ORIGIN
 Query Match 100.0%; Score 752; DB 6; Length 752;
 Best Local Similarity 100.0%; Pred. No. 1.6e-125;
 Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATCTCGTGGGCGCTGACCTTCTCTCTGAGAGCGCGGACAGGCTCCGAGCCATGACAG 60
 DB 1 ATCTCGTGGGCGCTGACCTTCTCTCTGAGAGCGCGGACAGGCTCCGAGCCATGACAG 60
 QY 61 CCGAAGCGCGGACAGGCGGCTTGAAGCGGCGATGCTGATGCGCCAGAGGCGCTTGACA 120
 DB 61 CCGAAGCGCGGACAGGCGGCTTGAAGCGGCGATGCTGATGCGCCAGAGGCGCTTGACA 120
 QY 121 TTCTGTATGGCCCAAGGCGGCGAATGCTGGCGGCGCCAGAGAGCGGCTGCCACGCGGCGCA 180
 DB 121 TTCTGTATGGCCCAAGGCGGCGAATGCTGGCGGCGCCAGAGAGCGGCTGCCACGCGGCGCA 180
 QY 121 TTCTGTATGGCCCAAGGCGGCGAATGCTGGCGGCGCCAGAGAGCGGCTGCCACGCGGCGCA 180
 DB 121 TTCTGTATGGCCCAAGGCGGCGAATGCTGGCGGCGCCAGAGAGCGGCTGCCACGCGGCGCA 180
 QY 181 GAGGTCCCGGCGGCGCAAGGCGGCGCTCGGCGCGGAGAGAGCGCGCCCGCGG 240
 DB 181 GAGGTCCCGGCGGCGCAAGGCGGCGCTCGGCGCGGAGAGAGCGCGCCCGCGG 240
 QY 241 GTCCGATGGCGGCGGCGCTTCAAGGCTGATGATGCTGCAATGCGGCGCGGCGGCGGCG 300
 DB 241 GTCCGATGGCGGCGGCGCTTCAAGGCTGATGATGCTGCAATGCGGCGCGGCGGCGGCG 300
 QY 301 CGAGAGCGCGCTGCTGATGCTGATGCTGCTGCGGAGACCCATGGAAGAGAG 360
 DB 301 CGAGAGCGCGCTGCTGATGCTGATGCTGCTGCGGAGACCCATGGAAGAGAG 360
 QY 361 AGCTGGCGCGGAGGCTTGCAGGATGCCCAAGGCTTCCGTCGCAAGGCGGCTTTC 420
 DB 361 AGCTGGCGCGGAGGCTTGCAGGATGCCCAAGGCTTCCGTCGCAAGGCGGCTTTC 420
 QY 421 TGAAGAGGTTCACTGTGTCCGCGCAACAATGACTGATGCTGATGCTGCGAGACAC 480
 DB 421 TGAAGAGGTTCACTGTGTCCGCGCAACAATGACTGATGCTGATGCTGCGAGACAC 480
 QY 481 GCCAAGTCAAGTCTCCATGAGCTCTGCTCCAGAGAGCTTCCCTGTGATGATGATCA 540
 DB 481 GCCAAGTCAAGTCTCCATGAGCTCTGCTCCAGAGAGCTTCCCTGTGATGATGATCA 540
 QY 541 GCGAGTCTTTTTCGCGCGCTTTTTCGCTCAAGCTCTCCAGAGAGAGGCGCTAAGGCC 600
 DB 541 GCGAGTCTTTTTCGCGCGCTTTTTCGCTCAAGCTCTCCAGAGAGAGGCGCTAAGGCC 600
 QY 601 AGCTGGCGCGCTTCTTAAGTCAATGCTCTCCAGAGAGAGGCGCTAAGGCC 660
 DB 601 AGCTGGCGCGCTTCTTAAGTCAATGCTCTCCAGAGAGAGGCGCTAAGGCC 660
 QY 661 GCCAGTCTTTTTCGCGCGCTTTTTCGCTCAAGCTCTCCAGAGAGAGGCGCTAAGGCC 720
 DB 661 GCCAGTCTTTTTCGCGCGCTTTTTCGCTCAAGCTCTCCAGAGAGAGGCGCTAAGGCC 720

FEATURES
source
Location/Qualifiers
1. 752
/organism="unknown"
BASE COUNT 126 a 230 c 256 g 140 t
ORIGIN

Query Match 100.0%; Score 752; DB 6; Length 752;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTCTGTGGGACCTGACCTTCTCTGAGAGCCGGGACAGAGGCTCCGAGCCATGACAG 60
DB 1 ATCTCTGTGGGACCTGACCTTCTCTGAGAGCCGGGACAGAGGCTCCGAGCCATGACAG 60
QY 61 CCGAAGGCCGGGGACACAGGGGGTTTCAGCGGGCGATGCTGATGGCCCAAGAGGCTCTGGCA 120
DB 61 CCGAAGGCCGGGGACACAGGGGGTTTCAGCGGGCGATGCTGATGGCCCAAGAGGCTCTGGCA 120
QY 121 TTCCTGATGGCCCAAGGGGGCAATGCTGAGCGGGCCCAAGAGAGCGGGTGCACAGGGCGCA 180
DB 121 TTCCTGATGGCCCAAGGGGGCAATGCTGAGCGGGCCCAAGAGAGCGGGTGCACAGGGCGCA 180
QY 181 GAGGTCCCGGGGGCGACAGGGGGCTTCGAGGGCCCGAGAGAGCGCCCGCGGG 240
DB 181 GAGGTCCCGGGGGCGACAGGGGGCTTCGAGGGCCCGAGAGAGCGCCCGCGGG 240
QY 241 GTCGGCATGGGGGGCGGGCTTCAGGGGCTGAAAGGATGCTGAGATGGGGGGCCAGGGGGC 300
DB 241 GTCGGCATGGGGGGCGGGCTTCAGGGGCTGAAAGGATGCTGAGATGGGGGGCCAGGGGGC 300
QY 301 CGAGAGAGCCGCTGCTTGAATCTTACCTCGCATGCTTCGCGACACCCATGAAAGCAG 360
DB 301 CGAGAGAGCCGCTGCTTGAATCTTACCTCGCATGCTTCGCGACACCCATGAAAGCAG 360
QY 361 AGCTGGCCCGGAGAGAGCTGGGCGGAGAGAGCCGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 361 AGCTGGCCCGGAGAGAGCTGGGCGGAGAGAGAGCCGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 TGAAGAGTTCACTGTGTCCGAGACATACTGACTATCCGACTGAGCTGAGAGACACC 480
DB 421 TGAAGAGTTCACTGTGTCCGAGACATACTGACTATCCGACTGAGCTGAGAGACACC 480
QY 481 GCCAATGCACTCTCTCATAGCTCCTGTCTCCAGAGAGCTTCCCTGTGATGTGATCA 540
DB 481 GCCAATGCACTCTCTCATAGCTCCTGTCTCCAGAGAGCTTCCCTGTGATGTGATCA 540
QY 541 CGCAGTGTCTTGCCTCGGTGTTTGTGAGCTCAGGCTCCCTCAGGGCAGAGGCGCTAAGCC 600
DB 541 CGCAGTGTCTTGCCTCGGTGTTTGTGAGCTCAGGCTCCCTCAGGGCAGAGGCGCTAAGCC 600
QY 601 AGCTGGCCCGCTTCTCTAGGTATGCTCTCTCCCTTGAAGAAATGGTCCACAGCAGTG 660
DB 601 AGCTGGCCCGCTTCTCTAGGTATGCTCTCTCCCTTGAAGAAATGGTCCACAGCAGTG 660
QY 661 GCCAGTTCAATGTGGGGGCGCTGATGTTTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 661 GCCAGTTCAATGTGGGGGCGCTGATGTTTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 TTCTGTGAAATAATAAATGAGCTACGAAAAA 752
DB 721 TTCTGTGAAATAATAAATGAGCTACGAAAAA 752

RESULT 2
AR159652 752 bp DNA linear PART 17-OCT-2001
LOCUS AR159652
DEFINITION Sequence 1 from patent US 6251603.
ACCESSION AR159652
VERSION AR159652.1 GI:16222376
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.

REFERENCE 1 (bases 1 to 752)
Jager, E., Stockert, E., Old, L. J., and Knuth, A.
METHOD for determining status of a cancerous condition by
determining antibodies to NY-ESO-1 in a patient sample
Patent: US 6251603-A 1 26-JUN-2001;
JOURNAL
TITLE
AUTHORS
JOURNAL
FEATURES
source
Location/Qualifiers
1. 752
/organism="unknown"
BASE COUNT 126 a 230 c 256 g 140 t
ORIGIN

Query Match 100.0%; Score 752; DB 6; Length 752;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTCTGTGGGACCTGACCTTCTCTGAGAGCCGGGACAGAGGCTCCGAGCCATGACAG 60
DB 1 ATCTCTGTGGGACCTGACCTTCTCTGAGAGCCGGGACAGAGGCTCCGAGCCATGACAG 60
QY 61 CCGAAGGCCGGGGACACAGGGGGTTTCAGCGGGCGATGCTGATGGCCCAAGAGGCTCTGGCA 120
DB 61 CCGAAGGCCGGGGACACAGGGGGTTTCAGCGGGCGATGCTGATGGCCCAAGAGGCTCTGGCA 120
QY 121 TTCCTGATGGCCCAAGGGGGCAATGCTGAGCGGGCCCAAGAGAGCGGGTGCACAGGGCGCA 180
DB 121 TTCCTGATGGCCCAAGGGGGCAATGCTGAGCGGGCCCAAGAGAGCGGGTGCACAGGGCGCA 180
QY 181 GAGGTCCCGGGGGCGACAGGGGGCTTCGAGGGCCCGAGAGAGCGCCCGCGGG 240
DB 181 GAGGTCCCGGGGGCGACAGGGGGCTTCGAGGGCCCGAGAGAGCGCCCGCGGG 240
QY 241 GTCGGCATGGGGGGCGGGCTTCAGGGGCTGAAAGGATGCTGAGATGGGGGGCCAGGGGGC 300
DB 241 GTCGGCATGGGGGGCGGGCTTCAGGGGCTGAAAGGATGCTGAGATGGGGGGCCAGGGGGC 300
QY 301 CGAGAGAGCCGCTGCTTGAATCTTACCTCGCATGCTTCGCGACACCCATGAAAGCAG 360
DB 301 CGAGAGAGCCGCTGCTTGAATCTTACCTCGCATGCTTCGCGACACCCATGAAAGCAG 360
QY 361 AGCTGGCCCGGAGAGAGCTGGGCGGAGAGAGCCGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 361 AGCTGGCCCGGAGAGAGCTGGGCGGAGAGAGAGCCGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 TGAAGAGTTCACTGTGTCCGAGACATACTGACTATCCGACTGAGCTGAGAGACACC 480
DB 421 TGAAGAGTTCACTGTGTCCGAGACATACTGACTATCCGACTGAGCTGAGAGACACC 480
QY 481 GCCAATGCACTCTCTCATAGCTCCTGTCTCCAGAGAGCTTCCCTGTGATGTGATCA 540
DB 481 GCCAATGCACTCTCTCATAGCTCCTGTCTCCAGAGAGCTTCCCTGTGATGTGATCA 540
QY 541 CGCAGTGTCTTGCCTCGGTGTTTGTGAGCTCAGGCTCCCTCAGGGCAGAGGCGCTAAGCC 600
DB 541 CGCAGTGTCTTGCCTCGGTGTTTGTGAGCTCAGGCTCCCTCAGGGCAGAGGCGCTAAGCC 600
QY 601 AGCTGGCCCGCTTCTCTAGGTATGCTCTCTCCCTTGAAGAAATGGTCCACAGCAGTG 660
DB 601 AGCTGGCCCGCTTCTCTAGGTATGCTCTCTCCCTTGAAGAAATGGTCCACAGCAGTG 660
QY 661 GCCAGTTCAATGTGGGGGCGCTGATGTTTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 661 GCCAGTTCAATGTGGGGGCGCTGATGTTTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 TTCTGTGAAATAATAAATGAGCTACGAAAAA 752
DB 721 TTCTGTGAAATAATAAATGAGCTACGAAAAA 752

RESULT 3
AR161582 752 bp DNA linear PART 17-OCT-2001
LOCUS AR161582
DEFINITION Sequence 1 from patent US 6255470.
ACCESSION AR161582

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2004, 01:46:32 ; Search time 3041 Seconds

(without alignments)
10116.422 Million cell updates/sec

Title: US-10-023-182-1

Perfect score: 752
Sequence: 1 ATCCCTGGGCGCCTGACCT.....TAAACTGAGCTACGAAAAA 752

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sbs:*

12: gb_sy:*

13: gb_un:*

14: gb_vt:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sbs:*

28: em_un:*

29: em_vt:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rnd:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	752	100.0	752	AR037873	AR037873 Sequence
2	752	100.0	752	AR159652	AR159652 Sequence
3	752	100.0	752	AR161582	AR161582 Sequence
4	752	100.0	752	AR164507	AR164507 Sequence
5	752	100.0	752	AX024691	AX024691 Sequence
6	752	100.0	752	AX024693	AX024693 Sequence
7	752	100.0	752	AX079857	AX079857 Sequence
8	752	100.0	752	AX107123	AX107123 Sequence
9	752	100.0	752	AX235020	AX235020 Sequence
10	752	100.0	752	AX558285	AX558285 Sequence
11	752	100.0	752	BD069515	BD069515 Isolated
12	752	100.0	752	HSU87459	U87459 Human autoi
13	750	99.7	806	AF038567	AF038567 Homo sapi
14	749	99.6	755	AR042537	AR042537 Sequence
15	749	99.6	755	HSAN03149	AR003149 Homo sapi
16	748.4	99.5	805	AX204870	AX204870 Sequence
17	681.6	90.6	767	AX024687	AX024687 Sequence
18	681.6	90.6	767	HSAN012834	AU012834 Homo sapi
19	680.6	90.5	755	AR042536	AR042536 Sequence
20	680.6	90.5	755	BD176600	BD176600 Method of
21	680.6	90.5	755	HSAN23041	AU223041 Homo sapi
22	599.2	79.7	679	AX024685	AX024685 Sequence
23	599.2	79.7	679	HSAN012833	AU012833 Homo sapi
24	540.2	70.1	1191	AX268307	AX268307 Sequence
25	527.4	70.1	540	AX204871	AX204871 Sequence
26	442.6	58.9	993	AX024689	AX024689 Sequence
27	442.6	58.9	993	HSAN012835	AU012835 Homo sapi
28	438.6	58.3	1002	AR042535	AR042535 Sequence
29	438.6	58.3	1002	HSAN23040	AU223040 Homo sapi
30	413.4	55.0	993	AX576894	AX576894 Sequence
31	413.4	55.0	993	BC002833	BC002833 Homo sapi
32	372	49.5	873	AX235028	AX235028 Sequence
33	372	49.5	873	HSAN275978	AU275978 Homo sapi
34	323.4	43.0	1751	AX235029	AX235029 Sequence
35	323.4	43.0	2630	AX235030	AX235030 Sequence
36	323.4	43.0	2630	HSAN275977	AU275977 Homo sapi
37	323.4	43.0	137217	HSAN211110	AU672302 Homo sapi
38	323.4	43.0	147234	AF277315	AF277315 Homo sapi
39	323.4	43.0	147234	AF277315	AF277315 Homo sapi
40	301.4	40.1	3245	HO6223093	AU23033 Homo sapi
41	301.4	40.1	142017	HSAN196H18	AL596249 Homo sapi
42	293.4	39.0	129083	AC144385	AC144385 Pan trogl
43	293.4	39.0	129083	AC144385	AC144385 Pan trogl
44	292.4	38.9	160656	AL592072	AL592072 Homo sapi
45	215.8	28.7	5572	AX251161	AX251161 Sequence

ALIGNMENTS

RESULT 1

AR037873

LOCUS AR037873 752 bp DNA linear PAT 29-SEP-1999

DEFINITION Sequence 1 from patent US 5804381.

ACCESSION AR037873

VERSION AR037873.1 GI:5956590

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 752)

AUTHORS Chen,Y.-T., Scanlan,M., Gure,A. and Old,L.J.

TITLE Isolated nucleic acid molecule encoding an esophageal cancer associated antigen, the antigen itself, and uses thereof

JOURNAL Patent: US 5804381-A 1 08-SEP-1999;

Feb 13 14:37:35 2004

us-10-023-182-1.mf

Page 11

.QY 102 GGCCCAAGAGGCGCTTCGATTCGTATGCGCCAGGGGCAATGCTGGGCGCCAGAGAG 161
 Db 561 GGCGCGGAGCGGCG 620
 QY 162 GCGGGATGCCAGCGGACGAGAGGTCCCGCGGCGCGAGGGGACAGCAAGAGGCGCTCGGAGCG 221
 Db 621 GATGGCG 674
 QY 222 GAGAGAGGCG 281
 Db 675 GCGCGGTGGCG 734
 QY 282 AGATGCGGGGCGAGGGGCGCGGAGAGCGCGCTGCTTGAATTCTACCTCGCGCATGCGCTTTC 341
 Db 735 GCGCACCTTCACCGCGCATATGCAACGCGTGGGCGCACCTGTGCACCGCGCGCGCGCGCG 794
 QY 342 GCGACACCCATGAGAGCAGAGCTGGCGCGCGAGAGCGTGGCGCGAGATGGCCCAACCGTT 401
 Db 795 GCATATGAATATGCGCTTCGGGCTTCGCGCACCGCGGCGTGTGGCGGCGGCGCGCACAC 854
 QY 402 CCGGTGCGAGGG 414
 Db 855 GGCGCGGAGCGG 867

Search completed: February 13, 2004, 13:20:28
Job time : 86 secs

Sequence 1, Application US/08791495
Patent No. 5811519
GENERAL INFORMATION:
APPLICANT: Leth, Bernard
APPLICANT: Lucas, Sophie
APPLICANT: De Smet, Charles
APPLICANT: Godelaine, Daniele
APPLICANT: Boon-Falleur, Thierry
TITLE OF INVENTION: LI-1 TUMOR SPECIFIC GENES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/791,495
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Van Amsterdam, John R.
REGISTRATION NUMBER: 40,212
REFERENCE/DOCKET NUMBER: L0461/7005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-2441
TELEFAX: 617-720-3500
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 217 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-791-495-1

Query Match 11.1%; Score 83.4; DB 1; Length 217;
Best Local Similarity 98.8%; Pred. No. 2.5e-10;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 455 TATTCGACTGACTCTGCAAGCACCGCCCAACTGCACTCTCATGCTCTGTTCCCA 514
DB 133 TAGCCGACTGACTCTGCTCAAGCACCGCCCAACTGCACTCTCATGCTCTGTTCCCA 192

QY 515 GCAGCTTCCCTGTTGATGTGATGC 539
DB 193 GCAGCTTCCCTGTTGATGTGATGC 217

RESULT 12
US-08-458-568A-11/c
Sequence 11, Application US/08458568A
Patent No. 5821339
GENERAL INFORMATION:
APPLICANT: Schaffer, Priscilla A.
APPLICANT: Yeh, Lily
TITLE OF INVENTION: Compositions and Methods for Treatment of Herpesvirus
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5821339r1s
STREET: One Liberty Place, 46th floor

CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,568A
FILING DATE: 02-JUNE-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/065,146
FILING DATE: 05-MAY-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Leary Ph.D., Kathryn R.
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: DFCI-0029
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 12001 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Herpes simplex virus
STRAIN: Herpes Simplex Virus Type 1
US-08-458-568A-11

Query Match 8.2%; Score 61.8; DB 1; Length 12001;
Best Local Similarity 50.4%; Pred. No. 3.9e-05;
Matches 177; Conservative 0; Mismatches 172; Indels 2; Gaps 1;

QY 25 TCTGAGAGCCCGGCGGAGAGGCTTCGAGCCATGCAAGCCGGAAGCCGGGCAAGGGGTT 84
DB 1505 TCCGGGCGGCGGCTGCGGAG 1446

QY 85 GCAGCGGCGATGCTGATGCGCCAGAGAGCCCTTGCACTTCTGATGCGCCAGGGGCGCATG 144
DB 1445 GAGCGGGGAGAGAGCGGGGAGAGAGCGGGGAGAGAGCGGGGAGAGAGCGGGGAGAG 1386

QY 145 CTGGGCGCCAGAGAGAGCGGGTGCACGGGCGGCGAGAGTCCCGGGGCGCAAGGGGAG 204
DB 1385 GAGCGGGGAGAGAGCGGGGAGAGAGCGGGGAGAGAGCGGGGAGAGAGCGGGGAGAG 1326

QY 205 CAAGGCGCTCGGGGCGGGGAGAGAGCGCCCGCGGGGTCCGATGCGGGGCGGGCTTAG 264
DB 1325 GAGCGGGGAGAGAGCGGGGAGAGAGCGGGGAGAGAGCGGGGAGAGAGCGGGGAGAG 1266

QY 265 GCGTAATAGA--TGCTGCAATGCGGGGCGAGGGGCGGAGAGCGCGCTGTGAGTT 322
DB 1265 GAGCGGGGAGAGAGCGGGGAGAGAGCGGGGAGAGAGCGGGGAGAGAGCGGGGAGAG 1206

QY 323 CTAAGTGGCAATGCTTTCGAGACACCCATGAGAGAGAGTGGCCCGCAG 373
DB 1205 CACACCCCCCGGGGAGTCCGCGCGGCGCTTTAAAGCGGTGGCGGCGGGGAG 1155

RESULT 13
US-09-103-840A-1
Sequence 1, Application US/09103840A
Patent No. 6284328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.

Db 481 CCAATGCAAGCTCTCAATCAGCTCTGTCTCCAGAGAGCTTCCCTGTGATGTGATAC 540
 Qy 542 GCAGTGTCTTCTGCCCCGTGTTTTTGGCTCAGCTCCCTCAGGCGAGAGCGCTTAAGCCCA 601
 Db 541 GCAGTGTCTTCTGCCCCGTGTTTTTGGCTCAGGCTCCCTCAGGCGAGAGCGCTTAAGCCCA 600
 Qy 602 GCCTGCGCGCCCTTCTAGGTGATGCTCTCCCTCAGGAGATGATCCAGACAGATG 661
 Db 601 GCCTGCGCGCCCTTCTAGGTGATGCTCTCCCTCAGGAGATGATCCAGACAGATG 660
 Qy 662 CCAATGCTATGTGAGGCGCTGATGTGTGCTGAGAGAGAGCGCTTACATGTTGTT 721
 Db 661 CCAATGCTATGTGAGGCGCTGATGTGTGCTGAGAGAGAGCGCTTACATGTTGTT 720
 Qy 722 TCTGTAGAAATTAAGCTGAGCTAGCAAAA 752
 Db 721 TCTGTAGAAATTAAGCTGAGCTAGCAAAA 751

RESULT 10

US-08-791-495-4
 Sequence 4, Application US/08791495
 Patent No. 5811519

GENERAL INFORMATION:

APPLICANT: Leith, Bernard
 APPLICANT: Lucas, Sophie
 APPLICANT: De Smet, Charles
 APPLICANT: Godelaine, Daniele
 APPLICANT: Boon-Palleu, Thierry
 TITLE OF INVENTION: LI-1 TUMOR SPECIFIC GENES
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/791,495
 FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: Van Amsterdam, John R.
 REGISTRATION NUMBER: 40,212
 REFERENCE/DOCKET NUMBER: 10461/7005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1002 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 65..697
 US-08-791-495-4

Query Match 58.3%; Score 438.6; DB 1; Length 1002;
 Best Local Similarity 72.1%; Pred. No. 3,6e-90;

Matches 704; Conservative 0; Mismatches 44; Indels 229; Gaps 1;

Qy 1 ATCCGTGTGGGCGCTGACCTTCTCTGTAGAGCCGGGAGAGAGGCTCCGAGCATGACAG 60
 Db 12 ATCTGTGTGGGCGCTGACCTTCTCTGTAGAGCCGGGAGAGAGGCTCCGAGCATGACAG 71
 Qy 61 CCGAAGGCGGAGCACAAGGAGGTTGACCGGCGATGCTGATGAGCCGAGAGGCTCTGCA 120
 Db 72 CCGAAGGCGGAGCACAAGGAGGTTGACCGGCGATGCTGATGAGCCGAGAGGCTCTGCA 131
 Qy 121 TTCCTATGAGCCCAAGGAGGCAATGCTGAGGCGCCAGAGAGGCGGAGTGCAGAGGAGCA 180
 Db 132 TTCCTATGAGCCCAAGGAGGCAATGCTGAGGCGCCAGAGAGGCGGAGTGCAGAGGAGCA 191
 Qy 181 GAGGTCCCGGAGCGCAGAGGAGCAGAGAGGCTCGGAGCGGAGAGAGGCGCCGCGAG 240
 Db 192 GAGGTCCCGGAGCGCAGAGGAGCAGAGAGGCTCGGAGCGGAGAGAGGCGCCGCGAG 251
 Qy 241 GTCCGATGAGCGGCGCGCTTCAAGGCTGAATGATGCTGAGATGCGGAGCGCAAGGAGC 300
 Db 252 GTCCGATGAGCGGCGCGCTTCAAGGCTGAATGATGATGAGATGCGGAGCGCAAGGAGC 311
 Qy 301 CCGAGAGCGGCTGCTTGAATCTACCTGCGCATGCTTTCGCGACACCCATGAGAGCAG 360
 Db 312 CCGAGAGCGGCTGCTTGAATCTACCTGCGCATGCTTTCGCGATGCTTTCGCGATGAGAGCAG 371
 Qy 361 AGCTGCGCGGAGAGAGCTGAGCCAGAGATGCCCAAGCGCTTCCGTCGAGAGGAGCTTC 420
 Db 372 AGCTGCGCGGAGAGAGCTGAGCCAGAGATGCCCAAGCGCTTCCGTCGAGAGGAGCTTC 431
 Qy 421 TGAAGAGTCACTGTGTCGCGCAATGCTGCTAT----- 457
 Db 432 TGAAGAGTCACTGTGTCGCGCAATGCTGCTAT----- 457
 Qy 458 ----- 457
 Db 492 AAGCGCTGAGCGAGATGAGGAGTGTGAGGCTGAGAGTCCGCTCCGAGAGGAGC 551
 Qy 458 ----- 457
 Db 552 AGAAGCTAGAGATCTCAGAACACCCAAACAGAGTCTCAGAACAGACCTGTGAC 611
 Qy 458 ----- 457
 Db 612 CAGGCGCGCGCACCAGAGAGAGCCAGGAGATGAGTGTGAGAGTGTGCTTATAG 671
 Qy 458 ----- CCGACTGACTGTGAGAGACAGACCGCCCACTGACG 491
 Db 672 TGATGTTCTGTGCCCTCAGCATTTAGCGACTGCTGCAACACCGCCCACTGACG 731
 Qy 492 CTCTCATGAGCTCTGTCTCAGACAGCTTCCCTGTGATGATGATCAGACAGTCTTT 551
 Db 732 CTCTCATGAGCTCTGTCTCAGACAGCTTCCCTGTGATGATGATCAGACAGTCTTT 791
 Qy 552 CTGCGCGTGTGTTGCTGACGCTCCCTCAGAGGAGAGGCGCTAACCCAGCGCTGCGCC 611
 Db 792 CTGCGCGTGTGTTGCTGACGCTCCCTCAGAGGAGAGGCGCTAACCCAGCGCTGCGCC 851
 Qy 612 CCTTCCTAGATGATGCTCTCCCTCAGAGAGATGCTCCAGACAGAGTGTGCTTCAAT 671
 Db 852 CCTTCCTAGATGATGCTCTCCCTCAGAGAGATGCTCCAGACAGAGTGTGCTTCAAT 911
 Qy 672 GTGAGGAGCTGATGTTGTGCTGAGAGAGAGCGCTTACATGTTGTTTCTGTAGAA 731
 Db 912 GTGAGGAGCTGATGTTGTGCTGAGAGAGAGCGCTTACATGTTGTTTCTGTAGAA 971
 Qy 732 ATAAAGCTGAGCTACGA 748
 Db 972 ATAAAGCTGAGCTACGA 988

RESULT 11
 US-08-791-495-1

Query Match 99.6%; Score 749; DB 1; Length 755;
 Best Local Similarity 100.0%; Pred. No. 4,4e-160;
 Matches 749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CTGTCGCGCCCTTACCTTCTCTCTGAGAGCGCGGCGAGAGGCTCCGAGCGCATGCAAGCCG 63
 DB 1 CTGTCGCGCCCTTACCTTCTCTCTGAGAGCGCGGCGAGAGGCTCCGAGCGCATGCAAGCCG 60
 QY 64 AAGCGCGCGGCGAGAGGCGGCTTCAAGCGCGGCGATGCTGAGCGCGCGCGCGCGCGCGCATTC 123
 DB 61 AAGCGCGCGGCGAGAGGCGGCTTCAAGCGCGGCGATGCTGAGCGCGCGCGCGCGCGCGCATTC 120
 QY 124 CTGATGCGCGCGCGAGAGGCGGCGATGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 183
 DB 121 CTGATGCGCGCGCGAGAGGCGGCGATGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
 QY 184 GTCGCCGCGGCGAGAGGCGGCG 243
 DB 181 GTCGCCGCGGCGAGAGGCGGCG 240
 QY 244 CGCATGCGCGCGCGCGCGCGCTTCAAGCGCGGCTGATGCTGAGCGCGCGCGCGCGCGCGCG 303
 DB 241 CGCATGCGCGCGCGCGCGCTTCAAGCGCGGCTGATGCTGAGCGCGCGCGCGCGCGCGCGCG 300
 QY 304 AGAGCGCGCGCGCTGAGCTTCACTGCGCGATGCTGCGCGATGCTGCGCGCGCGCGCGCGCG 363
 DB 301 AGAGCGCGCGCTGAGCTTCACTGCGCGATGCTGCGCGATGCTGCGCGCGCGCGCGCGCGCG 360
 QY 364 TGGCCCGCGAGAGCG 423
 DB 361 TGGCCCGCGAGAGCG 420
 QY 424 AGAGCTTCACTGCTGCTGCGCGAGACATGCTGCTGCGCGATGCTGCGCGCGCGCGCGCG 483
 DB 421 AGAGCTTCACTGCTGCTGCGCGAGACATGCTGCTGCGCGATGCTGCGCGCGCGCGCGCG 480
 QY 484 AACTGAGCTCTCATGAGCTCTGCTGCTGCGAGAGCTTCCCTGATGATGATGATGATGATGAT 543
 DB 481 AACTGAGCTCTCATGAGCTCTGCTGCTGCGAGAGCTTCCCTGATGATGATGATGATGATGAT 540
 QY 544 AGGCTTTTCTGCGCGCGGTTTTTGGCTGAGCTCTGCGCGCGCGCGCGCGCGCGCGCGCG 603
 DB 541 AGGCTTTTCTGCGCGCGGTTTTTGGCTGAGCTCTGCGCGCGCGCGCGCGCGCGCGCGCG 600
 QY 604 CTGCGCGCGCTTCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 663
 DB 601 CTGCGCGCGCTTCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 QY 664 AGTTCATTTGCGCGCGCGCTGATTTGTTGCTGCTGAGAGAGAGCGGCTTACATGTTGTTTC 723
 DB 661 AGTTCATTTGCGCGCGCGCTGATTTGTTGCTGCTGAGAGAGAGCGGCTTACATGTTGTTTC 720
 QY 724 TGTAGAAAATTAAGTGAAGTGAAGAAA 752
 DB 721 TGTAGAAAATTAAGTGAAGTGAAGAAA 749

RESULT 9
 US-08-791-495-6
 / Sequence 6, Application US/08791495
 / Patent No. 5811519
 / GENERAL INFORMATION:
 / APPLICANT: Letch, Bernard
 / APPLICANT: Lucas, Sophie
 / APPLICANT: De Smet, Charles
 / APPLICANT: Godelaine, Daniele
 / APPLICANT: Boon-Falleur, Thierry
 / TITLE OF INVENTION: LL-1 TUMOR SPECIFIC GENES
 / NUMBER OF SEQUENCES: 14
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Molt, Greenfield & Sacks, P.C.
 / STREET: 600 Atlantic Avenue

CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02210
 COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/791,495
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Van Amsterdam, John R.
 REGISTRATION NUMBER: 40,212
 REFERENCE/DOCKET NUMBER: 10461/7005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 755 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 53..595
 US-08-791-495-6

Query Match 90.5%; Score 680.6; DB 1; Length 755;
 Best Local Similarity 94.1%; Pred. No. 1.1e-144;
 Matches 707; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 2 TCTCTGCGCGCGCGCGCGCTTCTCTCTGAGAGCGCGGCGAGAGGCTCCGAGCGCATGCGAGC 61
 DB 1 TCTCTGCGCGCGCGCGCGCTTCTCTCTGAGAGCGCGGCGAGAGGCTCCGAGCGCATGCGAGC 60
 QY 62 CGAAGCGCGGCGCAAGCGGCGTTGAGCGGCGGATGCTGATGCGCGCGAGAGGCGCGCGCAT 121
 DB 61 CGAAGCGCGGCGCAAGCGGCGTTGAGCGGCGGATGCTGATGCGCGCGAGAGGCGCGCGCAT 120
 QY 122 TCTGATGCGCGCGCGCGCGCATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 181
 DB 121 TCTGATGCGCGCGCGCGCGCATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
 QY 182 AGTCCCGCGGCG 241
 DB 181 AGTCCCGCGGCG 240
 QY 242 TCCGATGCGCGCGCGCGCGCTTGAAGGCTGATGCTGCGAGATGCTGCGAGATGCGCGCGCG 301
 DB 241 TCCGATGCGCGCGCGCGCGCTTGAAGGCTGATGCTGCGAGATGCTGCGAGATGCGCGCGCG 300
 QY 302 GGAAGCGCGCGCTGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
 DB 301 GGAAGCGCGCGCTGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 QY 362 GCTGCGCGCGCGAGAGCTGCG 421
 DB 361 GCTGCGCGCGCGAGAGCTGCG 420
 QY 422 GAAAGAGTTCATGCTGCTGCG 481
 DB 421 GAAAGAGTTCATGCTGCTGCG 480
 QY 482 CCAACTGCACTTCCATCACTGCTCTGCTTCCAGAGAGCTTCCCTGTTGATGATGATCAAC 541


```

QY 301 CGAGAGCGCCGCTTGAATTCACCTGCGCATGCTTTCCGCAACCCCATGGAAGCAG 360
DB 301 CGAGAGCGCCGCTTGAATTCACCTGCGCATGCTTTCCGCAACCCCATGGAAGCAG 360
QY 361 AGCTGGCCCGCAGAGAGCTGGCCAGAGATGCCACCGCTTCCCGTCCAGAGGGTCTTC 420
DB 361 AGCTGGCCCGCAGAGAGCTGGCCAGAGATGCCACCGCTTCCCGTCCAGAGGGTCTTC 420
QY 421 TGAAGAGATTCACCTGCTCCGCGCAACATCTGACTATCCGACATGATGTCAGACACAC 480
DB 421 TGAAGAGATTCACCTGCTCCGCGCAACATCTGACTATCCGACATGATGTCAGACACAC 480
QY 481 GCCAATGACATCTCCATCAGCTCCCTGCTCCAGAGAGCTTCCCTGTTGATGTGATCA 540
DB 481 GCCAATGACATCTCCATCAGCTCCCTGCTCCAGAGAGCTTCCCTGTTGATGTGATCA 540
QY 541 CGCAGTGTCTTTCGCCCTGTTTGGCTCAGCTCCCTCAGGCGAGAGCGCTTAAGCCC 600
DB 541 CGCAGTGTCTTTCGCCCTGTTTGGCTCAGCTCCCTCAGGCGAGAGCGCTTAAGCCC 600
QY 601 AGCTGGCGCCCTTCTAGATGCTGCTCCCTCAGGAGATGCTCCAGACAGAGTG 660
DB 601 AGCTGGCGCCCTTCTAGATGCTGCTCCCTCAGGAGATGCTCCAGACAGAGTG 660
QY 661 GCCAGTTCATTTGGGGGCGCTGATTTGTTGTCGCTGAGAGAGAGCGCTTACATGTTGT 720
DB 661 GCCAGTTCATTTGGGGGCGCTGATTTGTTGTCGCTGAGAGAGAGCGCTTACATGTTGT 720
QY 721 TTCTGTAGAAAATAAACTAGCTACGAAAAA 752
DB 721 TTCTGTAGAAAATAAACTAGCTACGAAAAA 752

```

```

RESULT 6
US-08-937-263B-1
; Sequence 1, Application US/08937263B
; Patent No. 6274145
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng; Scanlan, Matthew;
; APPLICANT: Gure, Ali; Old, Lloyd J.; Jager, Elke;
; APPLICANT: Alexander, Knuth; Dzijhout, Jan W.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE
; TITLE OF INVENTION: ENCODING CANCER ASSOCIATED ANTIGEN, THE ANTIGEN
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski, L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/937,263B
; FILING DATE: September 15, 1997
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/725,182
; FILING DATE: October 3, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sinn, Eric, Patent Agent
; REGISTRATION NUMBER: 40,177
; REFERENCE/DOCKET NUMBER: LUD 5466.1
; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 752 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-937-263B-1

```

```

Query Match 100.0%; Score 752; DB 3; Length 752;
Best Local Similarity 100.0%; Pred. No. 9,4e-161;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 ATCTCTGTGGGCTTCTGACCTTCTCTCTGAGAGCCGGGCAAGAGGTCCGGAGCCATGACG 60
DB 1 ATCTCTGTGGGCTTCTGACCTTCTCTCTGAGAGCCGGGCAAGAGGTCCGGAGCCATGACG 60
QY 61 CCGAAGGCGGGGACACAGGGGGTTTCAGACGGGCGATGCTGATGGCCAGAGAGCCCTGGCA 120
DB 61 CCGAAGGCGGGGACACAGGGGGTTTCAGACGGGCGATGCTGATGGCCAGAGAGCCCTGGCA 120
QY 121 TTCTGTATGCTCCAGGGGGCAATGCTGCGGCGCCAGAGAGAGCGGTGCCACGGGCGGCA 180
DB 121 TTCTGTATGCTCCAGGGGGCAATGCTGCGGCGCCAGAGAGAGCGGTGCCACGGGCGGCA 180
QY 181 GAGGTCCCGGGGCGCAGAGGGGACAGAGAGGCTCGGGGCGGGAGAGAGCGCCCGGGG 240
DB 181 GAGGTCCCGGGGCGCAGAGGGGACAGAGAGGCTCGGGGCGGGAGAGAGCGCCCGGGG 240
QY 241 GTCCGATGCGGGGCGGCTTCAAGGCTGAATGATGCTGCAATGCGGGGCGGAGAGG 300
DB 241 GTCCGATGCGGGGCGGCTTCAAGGCTGAATGATGCTGCAATGCGGGGCGGAGAGG 300
QY 301 CGAGAGCGCGCTGCTGATGCTTACCTCGCATGCTTTGCGGACACCCATGGAAGCAG 360
DB 301 CGAGAGCGCGCTGCTGATGCTTACCTCGCATGCTTTGCGGACACCCATGGAAGCAG 360
QY 361 AGCTGGCCCGCAGAGAGCTGGCCAGAGATGCCACCGCTTCCCGTCCAGAGGGTCTTC 420
DB 361 AGCTGGCCCGCAGAGAGCTGGCCAGAGATGCCACCGCTTCCCGTCCAGAGGGTCTTC 420
QY 421 TGAAGAGTTCACCTGCTCCGCGCAACATCTGACTATCCGACATGCTGCGAGACAC 480
DB 421 TGAAGAGTTCACCTGCTCCGCGCAACATCTGACTATCCGACATGCTGCGAGACAC 480
QY 481 GCCAATGACATCTCCATCAGCTCCCTGCTCCAGAGAGCTTCCCTGTTGATGTGATCA 540
DB 481 GCCAATGACATCTCCATCAGCTCCCTGCTCCAGAGAGCTTCCCTGTTGATGTGATCA 540
QY 541 CGCAGTGTCTTTCGCCCTGTTTGGCTCAGCTCCCTCAGGCGAGAGCGCTTAAGCCC 600
DB 541 CGCAGTGTCTTTCGCCCTGTTTGGCTCAGCTCCCTCAGGCGAGAGCGCTTAAGCCC 600
QY 601 AGCTGGCGCCCTTCTAGATGCTGCTCCCTCAGGAGATGCTCCAGACAGAGTG 660
DB 601 AGCTGGCGCCCTTCTAGATGCTGCTCCCTCAGGAGATGCTCCAGACAGAGTG 660
QY 661 GCCAGTTCATTTGGGGGCGCTGATTTGTTGTCGCTGAGAGAGAGCGCTTACATGTTGT 720
DB 661 GCCAGTTCATTTGGGGGCGCTGATTTGTTGTCGCTGAGAGAGAGCGCTTACATGTTGT 720
QY 721 TTCTGTAGAAAATAAACTAGCTACGAAAAA 752
DB 721 TTCTGTAGAAAATAAACTAGCTACGAAAAA 752

```

```

RESULT 7
US-09-751-798-1
; Sequence 1, Application US/09751798
; Patent No. 6525177
; GENERAL INFORMATION:
; APPLICANT: Stockert, Elisabeth; Jager, Elke;
; APPLICANT: Chen, Yao-Tseng; Scanlan, Matthew;
; APPLICANT: Knuth, Alexander; Old, Lloyd J.
; TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer
; TITLE OF INVENTION: Associated Proteins, Uses Thereof,

```

APPLICATION NUMBER: 08/937,263
 FILING DATE: September 15, 1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/752,182
 FILING DATE: 03-October-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Hanson, No. 6252052man D.
 REGISTRATION NUMBER: 30,946
 REFERENCE/DOCKET NUMBER: LUD 5466.3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 688-9200
 TELEFAX: (212) 838-3884
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 752 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-09-062-422-1

Query Match 100.0%; Score 752; DB 3; Length 752;

Best Local Similarity 100.0%; Pred. No. 9,4e-161; Indels 0; Gaps 0;

Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATCTCTGAGGCGCTGACCTTCTCTCTGAGAGCCGCGGCAAGAGCTCCGAGCCATGCAAG 60
 1 ATCTCTGAGGCGCTGACCTTCTCTCTGAGAGCCGCGGCAAGAGCTCCGAGCCATGCAAG 60
 61 CCGAAGGCGGCGGCAAGAGGCGGTTGAGAGGCGGATGAGTGGCCAGAGGCGCTGGCA 120
 61 CCGAAGGCGGCGGCAAGAGGCGGTTGAGAGGCGGATGAGTGGCCAGAGGCGCTGGCA 120
 121 TTCTGTATGCGCCAGAGGCGCAATGCTGCGGCGCCAGAGAGCGGCGTCCACGCGCGCA 180
 121 TTCTGTATGCGCCAGAGGCGCAATGCTGCGGCGCCAGAGAGCGGCGTCCACGCGCGCA 180
 181 GAGGTCCCGGCGGCGCAAGAGGCGGCTCGGCGCGCGGAGAGAGCGCGCGCGG 240
 181 GAGGTCCCGGCGGCGCAAGAGGCGGCTCGGCGCGCGGAGAGAGCGCGCGCGG 240
 241 GTCCGATGCGCGCGCGCTTCAAGGCTGATGATGCTGCAAGTGCAGGCGCGCGCG 300
 241 GTCCGATGCGCGCGCGCTTCAAGGCTGATGATGCTGCAAGTGCAGGCGCGCGCG 300
 301 CCGAAGGCGGCGGCTTCTGAGTCTTACTTCCGATGCTTCCGCGCAACCCATGGAAGAG 360
 301 CCGAAGGCGGCGGCTTCTGAGTCTTACTTCCGATGCTTCCGCGCAACCCATGGAAGAG 360
 361 AGCTGCGCGCGAGAGAGCTGCGCGAGATGCCCAACCGCTTCCGTCGCAAGGCGTCTTC 420
 361 AGCTGCGCGCGAGAGAGCTGCGCGAGATGCCCAACCGCTTCCGTCGCAAGGCGTCTTC 420
 421 TGAAGAGTTCAGTGTGTCGCGGCAATCTGATCTGATGCTGATGCTGATGCTGATGCT 480
 421 TGAAGAGTTCAGTGTGTCGCGGCAATCTGATCTGATGCTGATGCTGATGCTGATGCT 480
 481 GCGAATGCGAGCTCTCATCATGCTCTGCTCTGAGAGCTTCCGTCGATGATGATGATGAT 540
 481 GCGAATGCGAGCTCTCATCATGCTCTGCTCTGAGAGCTTCCGTCGATGATGATGATGAT 540
 541 CCGAAGGCGGCGGCAAGAGGCGGTTGAGAGGCGGTTGAGAGGCGGTTGAGAGGCGGTT 600
 541 CCGAAGGCGGCGGCAAGAGGCGGTTGAGAGGCGGTTGAGAGGCGGTTGAGAGGCGGTT 600
 601 AGCTGCGCGCGCTTCTGAGTCTTACTTCCGATGCTTCCGTCGATGATGATGATGATGAT 660
 601 AGCTGCGCGCGCTTCTGAGTCTTACTTCCGATGCTTCCGTCGATGATGATGATGATGAT 660
 661 GCGAGTTCATGTTGCGGCGCGCTGATGTTGCTGCTGAGAGAGAGAGCGCTTACATGTTGT 720
 661 GCGAGTTCATGTTGCGGCGCGCTGATGTTGCTGCTGAGAGAGAGAGCGCTTACATGTTGT 720
 721 TTCTGTATGAGAAATTAAGTGAAGTGAAGAA 752

Db 721 TTCTGTATGAGAAATTAAGTGAAGTGAAGAA 752

RESULT 5

US-09-396-184-1
 Sequence 1, Application US/09396184
 Patent No. 6255470

GENERAL INFORMATION:
 APPLICANT: Yao-Tseung Chen, Scanlan, Matthew, Gure, Ali,
 Old, Lloyd

TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE ENCODING
 AN ESCAPEAL CANCER ASSOCIATED ANTIGEN, THE ANTIGEN ITSELF
 AND USES THEREOF

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Felle & Lynch
 STREET: 805 Third Avenue
 CITY: New York City
 STATE: New York
 COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage

COMPUTER: IBM

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/396,184
 FILING DATE: 14-Sep-1999
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/013,150
 FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
 NAME: Hanson, No. 6255470man D.
 REGISTRATION NUMBER: 30,946
 REFERENCE/DOCKET NUMBER: LUD 5466
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 688-9200
 TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 752 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-396-184-1

Query Match 100.0%; Score 752; DB 3; Length 752;

Best Local Similarity 100.0%; Pred. No. 9,4e-161;

Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATCTCTGAGGCGCTGACCTTCTCTCTGAGAGCCGCGGCAAGAGCTCCGAGCCATGCAAG 60
 1 ATCTCTGAGGCGCTGACCTTCTCTCTGAGAGCCGCGGCAAGAGCTCCGAGCCATGCAAG 60
 61 CCGAAGGCGGCGGCAAGAGGCGGTTGAGAGGCGGATGAGTGGCCAGAGGCGCTGGCA 120
 61 CCGAAGGCGGCGGCAAGAGGCGGTTGAGAGGCGGATGAGTGGCCAGAGGCGCTGGCA 120
 121 TTCTGTATGCGCCAGAGGCGCAATGCTGCGGCGCCAGAGAGCGGCGTCCACGCGCGCA 180
 121 TTCTGTATGCGCCAGAGGCGCAATGCTGCGGCGCCAGAGAGCGGCGTCCACGCGCGCA 180
 181 GAGGTCCCGGCGGCGCAAGAGGCGGCTCGGCGCGCGGAGAGAGCGCGCGCGG 240
 181 GAGGTCCCGGCGGCGCAAGAGGCGGCTCGGCGCGCGGAGAGAGCGCGCGCGG 240
 241 GTCCGATGCGCGCGCGCTTCAAGGCTGATGATGCTGCAAGTGCAGGCGCGCGCG 300
 241 GTCCGATGCGCGCGCGCTTCAAGGCTGATGATGCTGCAAGTGCAGGCGCGCGCG 300

QY 721 TTCTGTAGAAAATAAACTGAGCTACGAAAAA 752
Db 721 TTCTGTAGAAAATAAACTGAGCTACGAAAAA 752

RESULT 3
US-09-359-503-1
Sequence 1, Application US/09359503
Patent No. 6251603

GENERAL INFORMATION:

APPLICANT: Jager, Elke; Stockert, Elisabeth; Old, Lloyd J;
APPLICANT: Knuth, Alexander
TITLE OF INVENTION: Method for Determining Status of A
TITLE OF INVENTION: Cancerous Condition By Determining Antibodies
TITLE OF INVENTION: a Patient Sample
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pulbright & Jaworski, L.L.P.
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10103

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage

COMPUTER: IBM

OPERATING SYSTEM: PC-DOS

SOFTWARE: WordPerfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/359,503

FILING DATE: July 23, 1999

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/165,546

FILING DATE: October 2, 1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/062,422

FILING DATE: April 17, 1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/937,263

FILING DATE: September 15, 1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/752,182

FILING DATE: 03-October-1996

ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 6251603man D.

REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: LUD 5466.1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 318-3000

TELEFAX: (212) 752-5958

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 752 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-09-359-503-1

Query Match 100.0%; Score 752; DB 3; Length 752;

Best local Similarity 100.0%; Pred. No. 9,4e-161; Indels 0; Gaps 0;

Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTCTGTGGGCGCTGACCTTCTCTCTGTAGAGCCGGGCGAGAGGCTCCGGAGCCATGCAAG 60

Db 1 ATCTCTGTGGGCGCTGACCTTCTCTCTGTAGAGCCGGGCGAGAGGCTCCGGAGCCATGCAAG 60

QY 61 CCGAAGGCGCGGGGCGACAGGGGGTTTGAGAGGGGCGATGCTGATGGCCCGAGAGAGCCCTGCGCA 120

Db 61 CCGAAGGCGCGGGGCGACAGGGGGTTTGAGAGGGGCGATGCTGATGGCCCGAGAGAGCCCTGCGCA 120

QY 121 TTCTGTATGAGCCGAGGGGGCAATGCTGAGCGGCCCGAGAGAGGCGGGGTCCACGAGCGCGCA 180

Db 121 TTCTGTATGAGCCGAGGGGGCAATGCTGAGCGGCCCGAGAGAGGCGGGGTCCACGAGCGCGCA 180

Db 121 TTCTGTATGAGCCGAGGGGGCAATGCTGAGCGGCCCGAGAGAGGCGGGGTCCACGAGCGCGCA 180
QY 181 GAGGTCCCCGGGGGCGACAGGGGCGACAAAGGCGCTCGGGGCGGGAGAGAGCGCCCGCGGG 240
Db 181 GAGGTCCCCGGGGGCGACAGGGGCGACAAAGGCGCTCGGGGCGGGAGAGAGCGCCCGCGGG 240
QY 241 GTCCGATGAGGGGCGGGGCTTCAAGGCTGAATGATGCTGAGATGAGGGGGCGAGGGGGCG 300
Db 241 GTCCGATGAGGGGCGGGGCTTCAAGGCTGAATGATGCTGAGATGAGGGGGCGAGGGGGCG 300
QY 301 CGAAGAGCGCGCTGCTGATGATTAACCTCGCCATGCTTTTGGCGACACCCATGAGAGAG 360
Db 301 CGAAGAGCGCGCTGCTGATGATTAACCTCGCCATGCTTTTGGCGACACCCATGAGAGAG 360
QY 361 AGCTGGCCCGAGAGAGGCTTCAAGGCTGAGATGATGCTGAGATGAGGGGGCGAGGGGGCG 420
Db 361 AGCTGGCCCGAGAGAGGCTTCAAGGCTGAGATGATGCTGAGATGAGGGGGCGAGGGGGCG 420
QY 421 TGAAGAGATGATGCTGATGCTGCGGCAACATGATGATGCTGATGCTGATGCTGATGCTGAT 480
Db 421 TGAAGAGATGATGCTGATGCTGCGGCAACATGATGATGCTGATGCTGATGCTGATGCTGAT 480
QY 481 GCCAAGTGAAGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 481 GCCAAGTGAAGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 541 CGCAGTGTCTTGTGCGCGGCTTTTGGCTGAGCGCTCCCGCGAGAGAGGGGCGTAAAGCC 600
Db 541 CGCAGTGTCTTGTGCGCGGCTTTTGGCTGAGCGCTCCCGCGAGAGAGGGGCGTAAAGCC 600
QY 601 AGCCTGAGGCGCGCTTCTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 601 AGCCTGAGGCGCGCTTCTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 661 GCGAGTTCATTTGTGGGGGCTGATGTTTGTGCTGAGAGAGAGAGGCTTACATGTTGT 720
Db 661 GCGAGTTCATTTGTGGGGGCTGATGTTTGTGCTGAGAGAGAGAGGCTTACATGTTGT 720
QY 721 TTCTGTAGAAAATAAACTGAGCTACGAAAAA 752
Db 721 TTCTGTAGAAAATAAACTGAGCTACGAAAAA 752

RESULT 4

US-09-062-422-1
Sequence 1, Application US/09062422
Patent No. 6252052

GENERAL INFORMATION:

APPLICANT: Stockert, Elisabeth; Jager, Elke;
APPLICANT: Chen, Yao-Tseng; Scanlan, Matthew; Knuth, Alexander;
TITLE OF INVENTION: ANTIBODIES WHICH BIND TO NY-ESO-1
TITLE OF INVENTION: CANCER ASSOCIATED PROTEINS, USES THEREOF, TRUNCATED
TITLE OF INVENTION: FORMS OF NY-ESO-1, AND HLA BINDING PEPTIDES DERIVED
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felle & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage

COMPUTER: IBM

OPERATING SYSTEM: PC-DOS

SOFTWARE: WordPerfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/062,422

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:


```

QY 61 CCGAAGCCCGGCGACAGAGGCGGTTGCAAGCGGCGATGATGCGCCCAAGAGGCGCTGCGCA 120
DB 61 CCGAAGCCCGGCGACAGAGGCGGTTGCAAGCGGCGATGATGCGCCCAAGAGGCGCTGCGCA 120
QY 121 TTCTGATGAGCCCAAGAGGCGCAATGCTGCGCGCCCAAGAGAGGCGGATGCCAAGGCGCGCA 180
DB 121 TTCTGATGAGCCCAAGAGGCGCAATGCTGCGCGCCCAAGAGAGGCGGATGCCAAGGCGCGCA 180
QY 181 GAGGTCCCGCGGCGCGAGAGGCGCGCAAGAGGCGCTCGGAGGCGGAGAGAGGCGCGCGCGG 240
DB 181 GAGGTCCCGCGGCGCGAGAGGCGCGCAAGAGGCGCTCGGAGGCGGAGAGAGGCGCGCGCGG 240
QY 241 GTCCGCGATGCGCGCGCGGCTTCAAGGCGTGAATGAATGCTGCAAGATGCGCGCGCGCGCGG 300
DB 241 GTCCGCGATGCGCGCGCGGCTTCAAGGCGTGAATGAATGCTGCAAGATGCGCGCGCGCGG 300
QY 301 CGAGAGCGCGCGCTGCTGAGTTCTACCTGCGCATGCGCTTGGCGACACCCATGGAAGAG 360
DB 301 CGAGAGCGCGCGCTGCTGAGTTCTACCTGCGCATGCGCTTGGCGACACCCATGGAAGAG 360
QY 361 AGCTGCGCGCGAGAGCGCTGCGCGCGAGATGCCCAACCGCTTCCGTCGCGAGGCGTTC 420
DB 361 AGCTGCGCGCGAGAGCGCTGCGCGCGAGATGCCCAACCGCTTCCGTCGCGAGGCGTTC 420
QY 421 TGAAGAGTTCACTGTGTCCGCAACATCTGACTATCCGACTGACTGTGCGACAC 480
DB 421 TGAAGAGTTCACTGTGTCCGCAACATCTGACTATCCGACTGACTGTGCGACAC 480
QY 481 GCGAATGCACTCTCATAGCTCTGCTCAGCAGCTTCCGTTGATGATGATCA 540
DB 481 GCGAATGCACTCTCATAGCTCTGCTCAGCAGCTTCCGTTGATGATGATCA 540
QY 541 CGCAGTCTTTCGCGCGGTTTTCGCTCAGCTTCCCTCAGGCGAGAGCGCTAAGCC 600
DB 541 CGCAGTCTTTCGCGCGGTTTTCGCTCAGCTTCCCTCAGGCGAGAGCGCTAAGCC 600
QY 601 AGCTGCGCGCGCTTCTAGTATGCTCCTCCTCAGGAGATGATGCTCCAGACAGATG 660
DB 601 AGCTGCGCGCGCTTCTAGTATGCTCCTCCTCAGGAGATGATGCTCCAGACAGATG 660
QY 661 GCCAGTTCAATTGCGGCGCTGATTTGTTGCTGCTGAGAGAGAGCGCTTAATGTTGT 720
DB 661 GCCAGTTCAATTGCGGCGCTGATTTGTTGCTGCTGAGAGAGAGCGCTTAATGTTGT 720
QY 721 TTCTGAGAAATAAATAAGCTAGCAAAAA 752
DB 721 TTCTGAGAAATAAATAAGCTAGCAAAAA 752

```

```

RESULT 2
US-09-013-150-1
; Sequence 1, Application US/09013150
; Patent No. 6069233
; GENERAL INFORMATION:
; APPLICANT: Yao-Tseng Chen, Seanlan, Matthew, Gure, Ali,
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE ENCODING
; TITLE OF INVENTION: AN ESOPHAGAL CANCER ASSOCIATED ANTIGEN, THE ANTIGEN ITSELF,
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pelfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:

```

```

APPLICATION NUMBER: US/09/013,150
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/725,182
FILING DATE: 03-October-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6069233man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5466
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 752 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-013-150-1

```

```

Query Match 100.0%; Score 752; DB 3; Length 752;
Best Local Similarity 100.0%; Pred. No. 9,4e-161;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 ATCTCGTGGGCGCTTACCTCTCTGAGAGCGCGGCGAGAGGCTCGGAGCGCATGACAG 60
DB 1 ATCTCGTGGGCGCTTACCTCTCTGAGAGCGCGGCGAGAGGCTCGGAGCGCATGACAG 60
QY 61 CCGAAGCCCGGCGACAGGCGGTTGCAAGCGGCGATGCTGATGCGCCAGAGAGGCGCTTGCA 120
DB 61 CCGAAGCCCGGCGACAGGCGGTTGCAAGCGGCGATGCTGATGCGCCAGAGAGGCGCTTGCA 120
QY 121 TTCTGATGAGCCCAAGAGGCGCAATGCTGCGCGCCAGAGAGGCGGATGCCAAGGCGCGCA 180
DB 121 TTCTGATGAGCCCAAGAGGCGCAATGCTGCGCGCCAGAGAGGCGGATGCCAAGGCGCGCA 180
QY 181 GAGGTCCCGCGGCGCGAGGCGCGCAAGAGGCGCTCGGAGCGCGAGAGAGGCGCGCGG 240
DB 181 GAGGTCCCGCGGCGCGAGGCGCGCAAGAGGCGCTCGGAGCGCGAGAGAGGCGCGCGG 240
QY 241 GTCCGATGAGCGCGCGGCTTCAAGGCGTGAATGATGCTGCAATCGCGGCGCAAGGCGGCG 300
DB 241 GTCCGATGAGCGCGCGGCTTCAAGGCGTGAATGATGCTGCAATCGCGGCGCAAGGCGGCG 300
QY 301 CGAGAGCCCGCTGCTGAGTTCTACTCGCATGCTTTCGCGACACCCATGGAAGAG 360
DB 301 CGAGAGCCCGCTGCTGAGTTCTACTCGCATGCTTTCGCGACACCCATGGAAGAG 360
QY 361 AGCTGCGCGCGAGAGCGCTGCGCGCGAGATGCCCAACCGCTTCCGTCGCGAGGCGTTC 420
DB 361 AGCTGCGCGCGAGAGCGCTGCGCGCGAGATGCCCAACCGCTTCCGTCGCGAGGCGTTC 420
QY 421 TGAAGAGTTCACTGTGTCCGCAACATCTGACTATCCGACTGACTGTGCGACAC 480
DB 421 TGAAGAGTTCACTGTGTCCGCAACATCTGACTATCCGACTGACTGTGCGACAC 480
QY 481 GCGAATGCACTCTCATAGCTCTGCTCAGCAGCTTCCGTTGATGATGATCA 540
DB 481 GCGAATGCACTCTCATAGCTCTGCTCAGCAGCTTCCGTTGATGATGATCA 540
QY 541 CGCAGTCTTTCGCGCGGTTTTCGCTCAGCTTCCCTCAGGCGAGAGCGCTAAGCC 600
DB 541 CGCAGTCTTTCGCGCGGTTTTCGCTCAGCTTCCCTCAGGCGAGAGCGCTAAGCC 600
QY 601 AGCTGCGCGCGCTTCTAGTATGCTCCTCCTCAGGAGATGATGCTCCAGACAGATG 660
DB 601 AGCTGCGCGCGCTTCTAGTATGCTCCTCCTCAGGAGATGATGCTCCAGACAGATG 660
QY 661 GCCAGTTCAATTGCGGCGCTGATTTGTTGCTGCTGAGAGAGAGCGCTTAATGTTGT 720
DB 661 GCCAGTTCAATTGCGGCGCTGATTTGTTGCTGCTGAGAGAGAGCGCTTAATGTTGT 720

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2004, 10:54:53 ; Search time 68 Seconds

(without alignments)
4881.178 Million cell updates/sec

Title: US-10-023-182-1

Perfect score: 752
Sequence: 1 ATCTCGTGGGCGCTGACCT.....TAAACTGAGCTACGAAAA 752

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 569978 segs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seg: *
2: /cgn2_6/ptodata/1/ina/5B COMB.seg: *
3: /cgn2_6/ptodata/1/ina/6A COMB.seg: *
4: /cgn2_6/ptodata/1/ina/6B COMB.seg: *
5: /cgn2_6/ptodata/1/ina/PTCUS COMB.seg: *
6: /cgn2_6/ptodata/1/ina/backfile1.seg: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	752	100.0	752	1	US-08-725-182C-1
2	752	100.0	752	3	US-09-013-150-1
3	752	100.0	752	3	US-09-359-503-1
4	752	100.0	752	3	US-09-062-422-1
5	752	100.0	752	3	US-09-396-184-1
6	752	100.0	752	3	US-08-937-263B-1
7	752	100.0	752	4	US-09-751-798-1
8	749	99.6	752	4	US-08-791-495-6
9	680.6	90.5	755	1	US-08-791-495-6
10	438.6	58.3	1002	1	US-08-791-495-6
11	83.4	11.1	217	1	US-08-791-495-6
12	61.8	8.2	12001	1	US-08-458-568A-11
13	59.2	7.9	4411529	3	US-09-103-840A-1
14	57.6	7.7	4403765	3	US-09-103-840A-2
15	52.6	7.0	4524	3	US-08-845-998-7
16	52.6	7.0	4524	3	US-09-206-537-7
17	52.6	7.0	4524	3	US-09-430-854-7
18	51	6.8	320	3	US-08-165-264-7
19	51	6.8	4403765	3	US-09-103-840A-2
20	50.6	6.7	4411529	3	US-09-103-840A-1
21	50.4	6.7	7218	1	US-08-232-463-14
22	49.6	6.6	1926	4	US-09-249-585A-2
23	49.6	6.6	1926	4	US-09-410-399-3
24	49.6	6.6	1926	4	US-09-050-863-2
25	49.6	6.6	2580	4	US-09-359-081-2
26	49.6	6.6	5452	2	US-09-130-114-1
27	49.6	6.6	5452	2	US-09-130-114-1

28	49.6	6.6	8705	4	US-09-647-344A-14	Sequence 14, Appl
29	49.6	6.6	9600	3	US-08-910-647-1	Sequence 1, Appl
30	49.6	6.6	9600	4	US-09-620-925-1	Sequence 1, Appl
31	49.6	6.6	10596	1	US-07-884-811-15	Sequence 15, Appl
32	49.6	6.6	10596	1	US-07-884-811-15	Sequence 15, Appl
33	49.6	6.6	10596	1	US-08-087-783A-15	Sequence 15, Appl
34	49.6	6.6	10596	1	US-08-194-088B-15	Sequence 15, Appl
35	49.6	6.6	10596	2	US-08-194-087-15	Sequence 15, Appl
36	49.6	6.6	10596	5	PCT-US93-04648-15	Sequence 15, Appl
37	49.2	6.5	7218	1	US-08-232-463-13	Sequence 14, Appl
38	49.2	6.5	320	3	US-09-165-264-8	Sequence 14, Appl
39	49.2	6.5	319	3	US-09-128-155-16	Sequence 16, Appl
40	47.8	6.4	152331	3	US-09-165-264-11	Sequence 16, Appl
41	46.8	6.2	320	3	US-09-165-264-14	Sequence 14, Appl
42	46.8	6.2	320	3	US-09-165-264-14	Sequence 14, Appl
43	46.4	6.2	1550	2	US-08-609-443B-17	Sequence 17, Appl
44	46.4	6.2	1550	4	US-08-851-896-17	Sequence 17, Appl
45	46.2	6.1	318	3	US-09-165-264-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-725-182C-1
Sequence 1, Application US/08725182C
Patent No. 5804381
GENERAL INFORMATION:
APPLICANT: Yao-Tseng Chen, Scanlan, Matthew, Gure, Ali,
APPLICANT: Old, Lloyd
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE ENCODING
TITLE OF INVENTION: AN ESOPHAGEAL CANCER ASSOCIATED ANTIGEN, THE ANTIGEN ITSELF,
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSER: Felife & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/725,182C
FILING DATE: 03-October-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5804381man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5466
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 752 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-725-182C-1
Query Match 100.0%; Score 752; DB 1; Length 752;
Best Local Similarity 100.0%; Pred. No. 9.4e-161;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 ATCTCGTGGGCGCTGACCTCTCTCTGAGCGCGGCGAGGCTCCGAGCATGACG 60
DB 1 ATCTCGTGGGCGCTGACCTCTCTCTGAGCGCGGCGAGGCTCCGAGCATGACG 60

FT DOMAIN 175 1188 TRIPLE-HELICAL REGION.
 FT DOMAIN 1189 1214 NONHELICAL REGION (C-TERMINAL).
 FT SITE 741 743 CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 1089 1091 CELL ATTACHMENT SITE (POTENTIAL).
 FT CARBOHYD 1361 1361 N-LINKED (GLCNAC) (POTENTIAL).
 SQ SEQUENCE 1460 AA; 138762 MW; 58E5674D2B570697 CRC64;

Alignment Scores:

Pred. No.: 0.000308 Length: 1460
 Score: 180.50 Matches: 87
 Percent Similarity: 36.52% Conservative: 20
 Best Local Similarity: 29.69% Mismatches: 93
 Query Match: 12.91% Indels: 93
 DB: 1 Gaps: 13

US-10-023-182-1 (1-752) x CA11_CANFA (1-1460)

QY 683 TCAGGCCCCCAATGAACTGCGCACTCGCTGGAGCAATTCCTAGGGAGAGGCAAT 624
 DB 645 AlaglyProProglYgluaAa-----GlylysProgly 655
 QY 623 GACCTAGGAAGGGGCGCCAGCTGGGCTTAGCGCTCTGCTGAGGGAGGCTGAGCCAA 564
 DB 656 GluInglYglYalProglYaspleugly--AlaProglYProserglYalaArglyglu 674
 QY 563 AAA-----CACGGGCAAGAAAGCACTGCGTATCCATCAACAGGAAAGCTGAGAG 510
 DB 675 ArgglYpHeProglYgluaArglyValInglYPro-----ProglYProalaglYPro 692
 QY 509 ACAGAGCTGATGAGAGAGCTGCAAGTTGGCGGTGCTGACAGCTCAGTCGATAGTCAG 450
 DB 693 ArgglYalaInglY--AlaProglYasnapglYalaYsglyAspalaaglY----- 709
 QY 449 TATGTTCGGGACACAGTGAATCCTTCAGAGACCCCTGGCAAGGGAAGCGGTGGGC 390
 DB 710 -----AlaProglYalaProglYserInglY 719
 QY 389 ATCTCTGGG-----CAAGGCTCTGCGGGCCAGCTGCTGCTCATGGGTGTCG 342
 DB 719 laProglYleuInglYMetProglYgluaArglyAlaInglYleuPro-GlyProlys 738
 QY 341 GAAAGGATGGCGAGGTGAACCTCAAGCAGCG----- 309
 DB 739 GlyAspArglyAspalaaglYProlysglyAlaAspGlyserProglYlysAspGlyVal 758
 QY 308 -----GCTCTCGGGCCCCCTGGCCCCGCGCATCTGAGCATCCATTCAAGCC 264
 DB 759 ArgglYleuInglYProilleglyProProglYProilleglyAlaProglYasplys-Gl 778
 QY 263 TGAAGCCGCGCG-----CATGCGAAGCCCGCGGGCGCT----- 227
 DB 778 YgluaInglYProserglYProalaglYProthrglyAlaArglyAlaProglYaspar 798
 QY 226 -----CCTCGCGGGCCCCGAGGCCCTTGTGCGCCCTGCGGCCCGGGAGCTCT 180
 DB 798 gglYgluProglYProProglYProalaglYpHealaglYPro--ProglYAlaAspGl 817
 QY 179 GCCCGCCGTGGCAACCCCTCTCTCTGGCGG-----CCAGCATTCGCCCC 135
 DB 817 YgluProglYAlaYsglygluProglYaspalaaglYAlaYsglyAspalaaglYProPr 837
 QY 134 TGGGCGCATCAGAGATGCCAGGCGCTCTGGGCGA----- 101
 DB 837 oglyProalaglYProthrglyProProglYProilleglyAsnValglYAlaProglYPr 857
 QY 100 -----TCAGCATCGCCC----- 89
 DB 857 olysglyAlaArglyYserInglYProProglYAlaThrglyPheProglYAlaIngl 877
 QY 88 -----GTGGAACCCCTGTGCCCCGCGCTTGGCGCTGCGATGCTCCGAGGCTCTGCCCG 33
 DB 877 YArgValglYProProglYProserglYasnapalaaglYProProglYProProglYProal 897

QY 32 CTCTCAGAGAGAAAGTCAAGGCCCAAGAGAGA 2
 DB 897 aglyysgluInglYglYysglYAlaArgly 907

Search completed: February 10, 2004, 17:34:39
 Job time : 24.5 secs

RA Kerehaw J.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: NEMATODE CUTICLES ARE COMPOSED LARGELY OF COLLAGEN-LIKE
 CC PROTEINS. THE CUTICLE FUNCTIONS BOTH AS AN EXOSKELETON AND AS A
 CC BARRIER TO PROTECT THE WORM FROM ITS ENVIRONMENT.
 CC -1- SUBUNIT: COLLAGEN POLYPEPTIDE CHAINS ARE COMPLEXED WITHIN THE
 CC CUTICLE BY DISULFIDE BONDS AND OTHER TYPES OF COVALENT CROSS-
 CC LINKS.
 CC -1- SIMILARITY: BELONGS TO THE CUTICULAR COLLAGEN FAMILY.
 CC -1- CAUTION: Ref.3 sequence differs from that shown due to erroneous
 CC gene model prediction.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J01047; AAA27988.1; -
 DR EMBL: V00147; CAA23463.1; -
 DR EMBL: 274472; CAA98942.1; ALT_SEQ.
 DR PIR: A31219; A31219.
 DR WormPep: P23H12.4; CE05707.
 DR GO: GO:0005576; C:extracellular; NAS.
 DR GO: GO:0040002; P:cuticle biosynthesis (sensu Nematoda); NAS.
 DR InterPro: IPR002486; Col_cuticle_N.
 DR InterPro: IPR000087; Collagen.
 DR Pfam: PF01484; Col_cuticle_N; 1.
 DR Pfam: PF01391; Collagen; 3.
 DR Cuticle; Connective tissue; Repeat; Multigene family; Collagen;
 KW Signal.
 FT SIGNAL 1 37 POTENTIAL.
 FT CHAIN 38 296 CUTICLE COLLAGEN 1.
 FT DOMAIN 100 129 TRIPLE-HELICAL REGION.
 FT DOMAIN 148 174 TRIPLE-HELICAL REGION.
 FT DOMAIN 178 204 TRIPLE-HELICAL REGION.
 FT DOMAIN 213 278 TRIPLE-HELICAL REGION.
 FT CONFLICT 256 A -> V (IN REF. 3).
 SQ SEQUENCE 296 AA; 28574 MW; 843B169D2C4526A CRC64;
 Alignment Scores:
 Pred. No.: 0.000319 Length: 296
 Score: 161.00 Matches: 66
 Percent Similarity: 37.31% Conservative: 9
 Best Local Similarity: 32.84% Mismatches: 62
 Query Match: 12.95% Indels: 64
 Gaps: 11
 DB: 1
 US-10-023-182-1 (1-752) x C001_CABEL (1-296)
 QY 426 CCTTGAAGACACCCCTGGACGGAAGCGTGGGCACTCTGGG----- 382
 DB |||||
 DB 70 PPOFARAGASARATHTHATRGALATRYGLYGLYPROGLVALAAsPronAlaPro 89
 QY 381 -----CCAGGCTCCTGGGCGGCGAGCTGCTGCATGCATG 349
 DB |||||
 DB 90 AaNLuGLInCYsGLuGLYsCYsLeuPProGLYProPProGLYProAlaGLYAlaProGLY 109
 QY 348 GTGTGGCGAAGACATGGGAGAGTAACTCAAGACAGGGGCTCTCGG---GCCCGCTGG 292
 DB |||||
 DB 110 LysPProGLYLySPro-----GLYATrPProGLYAlaPProGLYThrPProGLYThrPProGL 127
 QY 291 CCGCGGATGTGACATCATTCAGCGCGCGAAGCGGCGCGCATGCGCGA-----CCCG 238
 DB |||||
 DB 127 LysPProPProVALAla-ProCYsGLuPProThrThrPProPProCYsLysPProCYsPProG 147
 QY 237 GCGGGGCGGCT---CCTCCGGGCGCC----- 215
 DB |||||
 DB 147 InGLYProPProGLYProPProGLYAlaPProGLYAsPProGLYGLuAlaGLY 167
 QY 214 -----GAGCGCTTGCTGCGCCCTGCGCGCGCGCGGAGACTCTGCG 176

DB 167 hPProGLYArGProGLYThrAsPAlaAlaPProGLYSePProGLYProArGLYProPProG 187
 QY 177 CGCGCGTGACACCGCC-----TTCCTGGGCGCGCGCATTG- 140
 DB |||||
 DB 187 LYProAlaGLYGLuAlaGLYAlaPProGLYProAlaGLYGLuPProGLYThrProAlaLies 207
 QY 139 -----CCCCCTGGGCGCATCAGAAATGCCAGGCGCTC 109
 DB 207 eRGLuPProLeuThrPProGLYAlaPProGLYGLuPProGLYAsPserLYProPProGLYProP 227
 QY 108 CTGGGCGCATCAGATGCGCGCTCGAA-----CCCCCTGTGCGCGCGCTTGCGGCT 58
 DB |||||
 DB 227 rGLYProPProGLYAlaPProGLYAsAAsPGLYProPProGLYProPProGLYProLYsGLY 247
 QY 57 GCATGCTCGGAGCGCTCGCGCGCTCTCAGAGAAAGTCAAG---GGCCACAGAGA 2
 DB |||||
 DB 247 laPProGLYProAsPGLYProPProGLYAla-----AsPGLYGLuSerGLYProPProGLY 264
 RESULT 15
 CALL_CANFA STANDARD; PRT: 1460 AA.
 ID CALL_CANFA
 AC Q9XST7;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Collagen alpha 1(I) chain precursor.
 GN COL1A1.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RA Campbell B.G., Mooton J.A.M., McLeod J.N., Minor R.R.;
 RT "Sequence of normal canine COL1A1 cDNA."
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: TYPE I COLLAGEN IS A MEMBER OF GROUP I COLLAGEN
 CC (FIBRILLAR FORMING COLLAGEN).
 CC -1- SUBUNIT: TRIMERS OF ONE ALPHA 2(I) AND TWO ALPHA 1(I) CHAINS.
 CC -1- PTM: Prolines at the third position of the tripeptide repeating
 CC unit (G-X-Y) are hydroxylated in some or all of the chains.
 CC -1- SIMILARITY: Contains 1 VWFC domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF153062; AAD34619.1; -
 DR InterPro: IPR000087; Collagen.
 DR InterPro: IPR000885; Fib_collagen_C.
 DR InterPro: IPR001007; VWFC_C.
 DR Pfam: PF01410; COLFI; 1.
 DR Pfam: PF01391; Collagen; 18.
 DR ProDom: PD000007; C1g_helix; 2.
 DR ProDom: PD002078; Fib_collagen_C; 1.
 DR SMART: SM00038; COLFI; 1.
 DR SMART: SM00214; VWFC; 1.
 DR PROSITE: PS01208; VWFC_1; 1.
 DR PROSITE: PS50184; VWFC_2; 1.
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 KW Glycoprotein; Collagen; Signal.
 FT SIGNAL 1 22
 FT PROPEP 23 157
 FT CHAIN 158 1214
 FT PROPEP 1215 1460
 FT DOMAIN 34 92
 FT DOMAIN 158 174
 NONHELICAL REGION (N-TERMINAL).

RN	Eur. J. Biochem.	59;113-118(1975).		
RP	SEQUENCE OF 416--445.			
RC	TISSUE=Skin:			
RX	MEDLINE=75008198; PubMed=4412529;			
RA	Fietzek P.P., Furchmayr H.; Kuehn K.;			
RT	"Comparative sequence studies on alpha2-CB2 from calf, human, rabbit			
RL	and pig-skin collagen."			
RN	Eur. J. Biochem.	47;257-261(1974).		
[5]				
RP	SEQUENCE OF 446-481.			
RC	TISSUE=Skin:			
RX	MEDLINE=75059250; PubMed=4435743;			
RA	Fietzek P.P., Kuehn K.;			
RT	"The covalent structure of collagen: amino acid sequence of the N-			
RL	terminal region of alpha2-CB3 from rat skin collagen and alpha2-CB3.5			
CC	from calf skin collagen."			
CC	Hoppe-Seyler's Z. Physiol. Chem.	355;647-650(1974).		
CC	-I- FUNCTION: TYPE I COLLAGEN IS A MEMBER OF GROUP I COLLAGEN			
CC	(FIBRILLAR FORMING COLLAGEN).			
CC	-I- SUBUNIT: TRIMERS OF ONE ALPHA 2(I) AND TWO ALPHA 1(I) CHAINS.			
CC	-I- TISSUE SPECIFICITY: FORMS THE FIBRILS OF TENDON, LIGAMENTS AND			
CC	BONES. IN BONES THE FIBRILS ARE MINERALIZED WITH CALCIUM			
CC	HYDROXYAPATITE.			
CC	-I- PM: Proline at the third position of the tripeptide repeating			
CC	unit (G-X-Y) are hydroxylated in some or all of the chains.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AB008683; BAA25171.1; -			
DR	InterPro; IPRO00087; Collagen_C.			
DR	InterPro; IPRO00885; Fib_collagen_C.			
DR	Pfam; PF01410; COLFI; 1.			
DR	Pfam; PF01391; Collagen; 18.			
DR	Pfdom; PD000007; Cig_helix; 3.			
DR	Pfdom; PD002078; Fib_collagen_C; 1.			
DR	SMART; SMO0038; COLFI; 1.			
KW	Extracellular matrix; Connective tissue; Repeat; Hydroxylation;			
KW	Collagen; Signal; Pyrrolidone carboxylic acid.			
FT	SIGNAL	1	24	POTENTIAL.
FT	PROPEP	25	79	AMINO-TERMINAL PROPEPTIDE.
FT	CHAIN	80	1100	COLLAGEN ALPHA 2(I) CHAIN.
FT	PROPEP	1101	1364	CARBOXY-TERMINAL PROPEPTIDE.
FT	MOD_RES	80	80	PYRROLIDONE CARBOXYLIC ACID.
FT	MOD_RES	84	84	CONVERTED TO AN ALDEHYDE GROUP THAT IS
FT				INVOLVED IN CROSS-LINKING (PROBABLE).
FT	MOD_RES	175	175	HYDROXYLATION (PROBABLE).
FT	MOD_RES	196	196	HYDROXYLATION (PROBABLE).
FT	MOD_RES	262	262	HYDROXYLATION (PROBABLE).
FT	MOD_RES	307	307	HYDROXYLATION (PROBABLE).
FT	MOD_RES	352	352	HYDROXYLATION (PROBABLE).
FT	CONFLICT	157	157	V -> P (IN REF. 3).
FT	CONFLICT	187	187	K -> T (IN REF. 3).
FT	CONFLICT	211	211	T -> K (IN REF. 3).
FT	CONFLICT	298	300	PGA -> AGP (IN REF. 3).
FT	CONFLICT	423	424	AT -> TA (IN REF. 4).
QO	SEQUENCE	1364 AA;	129064 MW;	5593F4DB95D119A CRC64;

Alignment Scores:	
Pred. No.:	0.0025
Length:	1364
Score:	187.00
Percent Similarity:	38.50%
Best local Similarity:	33.00%
Query Match:	13.02%
Ds:	1
US-10-023-182-1 (1-752) x CA21_BOVIN (1-1364)	
Length:	1364
Matches:	66
Conservative:	11
Mismatches:	53
Indels:	70
Gaps:	10

Oy	396	GTGGGGGCAATCCGGG-----	-----CCAGGCTCTGGGG	367
Dd	625	ValGlyAlaProGlyThrAlaGlyProSerGlyProSerGlyLeuProGlyValuArgGly	-----	644
Oy	366	CAAGCTTCGCTTCATGGGTGCGGAAAGCATGCGAGGTAGACTCAAG-----	-----	315
Dd	645	AlaAlaGly11LePro-GlyGlyLysGlyGluLysGlyGluThrGlyLeuArgGlyAsp11	-----	664
Oy	314	-----	CAAGCGGCTCTCGGCGCCCTGGCC	289
Dd	664	eglYserProGlyArgAspGlyAlaArgGlyAlaProGlyAlaAlaGlyAlaProGlyPro	-----	684
Oy	288	CGCATTCGACGATCCATTCATCAAGCTTGAAGCGGCGCG-----	-----CATGCG	244
Dd	684	oAlaGlyAlaAlaAnglyAspArg-GlyGluAlaGlyProAlaGlyProAlaGlyProAlaG	-----	704
Oy	243	GATCCCGCGGGGGCGCT-----	-----CTCCGGGCGCCGAGGCCCTTG	205
Dd	704	LyrProArgLysPheProGlyGluArgGlyGluValGlyProAlaGlyProAsnGlyPheA	-----	724
Oy	204	CTGCCCCCTGCG-----	-----	194
Dd	724	laGlyProAlaGlyAlaAlaGlyGlnProGlyAlaAlaGlyGluArgGlyThrLysGlyP	-----	744
Oy	193	-----	-----CCCGGGGACCTCTGCGCGCGCGCGGACCGGCTCTGCTGGGC	151
Dd	744	roLysGlyGluAsnGlyProValGlyProThrGlyProValGlyAlaAlaGlyProSerG	-----	764
Oy	150	CGCGACGATTCGCCCCCTGGGCGCATCAGAGATG-----	-----CGAGGCGTCTCTGGG	104
Dd	764	lyProAsnGlyProProGlyProAlaGlySerArgGlyAspGlyGlyProProGlyAlaT	-----	784
Oy	103	-----	-----CCATCAGATCGCCCGTCGAACCCCTCTGCGCCCGGCTTGGCGCTGCATGG	52
Dd	784	hriGlyPheProGlyAlaAlaGlyArgThrGlyPro--ProGlyProSerGlyTLeSerG	-----	803
Oy	51	CTCCGAGAGCCTCGCCGCGC--TCTCAGAGAGAGAGT--	-----CAGGGCGCCACGAGGA	2
Dd	803	lyProProGlyProProGlyProAlaGlyLysGluGlyLeuArgGlyProArgGly	-----	821

RESULT 14

CCOL	CABEL	STANDARD	PRT	296	AA.	
ID	CCOL	CABEL	STANDARD	PRT	296	AA.
AC	P08124	G19763				
DT	01-AUG-1988	(Rel. 08, Created)				
DT	01-AUG-1988	(Rel. 08, Last sequence update)				
DT	15-SEP-2003	(Rel. 42, Last annotation update)				
DE	Cuticle collagen 1 precursor (squt protein 3).					
GN	SQT-3 OR COL-1 OR DPY-15 OR F23H1.4.					
OS	Caenorhabditis elegans.					
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;					
OC	Rhabditidae; Peridderinae; Caenorhabditis.					
NCBI	Taxid=6239;					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=Bristol N2.					
RC	MEDLINE=8105075; PubMed=2578467;					
RA	Kramer J.M., Cox G.N., Hirsch D.;					
RT	"Expression of the Caenorhabditis elegans collagen genes col-1 and					
RT	col-2 is developmentally regulated."					
RU	J. Biol. Chem. 260:1945-1951(1985).					
RM	[3]					
RM	SEQUENCE FROM N.A.					
RC	STRAIN=Bristol N2;					
RC	MEDLINE=81050944; PubMed=7139711;					
RA	Kramer J.M., Cox G.N., Hirsch D.;					
RT	"Comparisons of the complete sequences of two collagen genes from					
RT	Caenorhabditis elegans."					
RU	Cell 30:559-606(1982).					
RM	[3]					
RM	SEQUENCE FROM N.A.					
RC	STRAIN=Bristol N2;					

```

RESULT 14
CCOL_CABEL
ID_CCOL_CABEL STANDARD; PRT; 296 AA.
AC P08124; 019763;
DT 01-AUG-1988 (rel. 08. Created)
DT 01-AUG-1988 (rel. 08. Last sequence update)
DT 15-SEP-2003 (rel. 42. Last annotation update)
DE Cuticle collagen 1 precursor (Squat protein 3) .
SQF-3 OR COL-1 OR DPY-15 OR F23H12.4.
GN Caenorhabditis elegans.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;
OC Rhabditidae; Ploloderinae; Caenorhabditis.
OX NCBI_TaxId=6239;
EN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2.
RX MEDLINE=85105075; PubMed=2578467;
RA Kramer J.M., Cox G.N., Hirsch D.;
RT "Expression of the Caenorhabditis elegans collagen genes col-1 and
   col-2 is developmentally regulated.";
RL J. Biol. Chem. 260:1945-1951 (1985).
EN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2.
RX MEDLINE=83050944; PubMed=7139711;
RA Kremer J.M., Cox G.N., Hirsch D.;
RT "Comparisons of the complete sequences of two collagen genes from
   Caenorhabditis elegans.";
RL Cell 30:599-606 (1982).
EN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;

```



```

QY 341 GAAAGCATGGCGAGGTAGACTCAAGACGCGC----- 309
DB 308 GYAAsparaglyAspAlaGlyProlylGlyAlaAspGlyAlaProglyLysAspGlyAla 327
QY 308 -----GCTCTCCGGCCCTGAGCCCGCGCATGTCAGCATCCATTGAGCC 264
DB 328 ArgGlyLeuThrGlyProIleGlyProProGlyProIleGlyAlaProGlyAspLys-GI 347
QY 263 TGAAGCGGCGCGC-----CCATGGGAGCCCGCGGCGCGCT----- 227
DB 347 yGIuAlaGlyProSerGlyProAlaGlyThrArgGlyAlaProGlyAspArgGlyLys 367
QY 226 ---CCTCCCGGCGCGGAGCCCTTGCCTGCGCCGCGGAGACCTTGCAGCCGCT 171
DB 367 oglyProProGlyProAlaGlyPheAlaGlyPro--ProGlyAlaAspGlyLysProGI 386
QY 170 GGCACCCGCGCTCTCTCGGCGC-----CCAGCATGCGCCCTGCGGCGCATC 126
DB 386 yAlaLysGlyGluProGlyAspAlaGlyAlaLysGlyAspAlaGlyProProGlyProAl 406
QY 125 AGGATGCGCAGGCGCTCTCGGCGC----- 101
DB 406 aglyProAlaGlyProProGlyProIleGlyAsnValGlyAlaProGlyProLysGlyAl 426
QY 100 ---TCAGCATCGCC-----GTGCA 84
DB 426 aaArgGlySerAlaGlyProProGlyAlaThrGlyPheProGlyAlaAlaGlyArgValGI 446
QY 83 ACCCGCTGCGCCGCGCTCGGCGCTGCGCATGCGCTGCGGAGCCTTGCAGGCTCAAG 24
DB 446 yProProGlyProSerGlyAsnAlaGlyProProGlyProProGlyProAlaGlyLysGI 466
QY 23 AGAAGTCAGGCGCCGCGAGCA 2
DB 466 uGlySerLysGlyProArgly 473

RESULT 10
CA21_CANFA STANDARD; PRT; 1366 AA.
AC 046352;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Collagen alpha 2(I) chain precursor.
GN COL1A2.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Campbell B.G., Wootton J.A.M., McLeod J.N., Minor R.R.;
RI Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: TYPE I COLLAGEN IS A MEMBER OF GROUP I COLLAGEN
CC (FIBRILLAR FORMING COLLAGEN).
CC -1- SUBUNIT: TRIMERS OF ONE ALPHA 2(I) AND TWO ALPHA 1(I) CHAINS.
CC -1- TISSUE SPECIFICITY: FORMS THE FIBRILS OF TENDON, LIGAMENTS AND
CC BONES. IN BONES THE FIBRILS ARE MINERALIZED WITH CALCIUM
CC HYDROXYAPATITE.
CC -1- PTM: Prolines at the third position of the tripeptide repeating
CC unit (G-X-Y) are hydroxylated in some or all of the chains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF035120; AAC64485.1; -.
CC InterPro; IPR000087; Collagen.

```

```

DR InterPro; IPR000885; Fib_collagen_C.
DR Pfam; PF01410; COLF; 1.
DR DR Pfam; PF01381; Collagen; 17.
DR ProDom; PD000007; C1G_helix; 3.
DR ProDom; PD002078; Fib_collagen_C; 1.
DR SMART; SM00038; COLF; 1.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Collagen; Signal; Pyroglutamate; Pyroglutamate; Pyroglutamate;
FT SIGNAL 1 24
FT PROPEP 25 79
FT CHAIN 80 1102
FT PROPEP 1103 1366
FT MOD_RES 80 80
FT MOD_RES 84 84
FT MOD_RES 84 84
FT INVOLVED IN CROSS-LINKING (BY
FT INVOLVED IN CROSS-LINKING (BY
SQ SEQUENCE 1366 AA; 129400 MW; CD936969E080BCD4 CRC64;

Alignment Scores:
Pred. No.: 0.000189 Length: 1366
Score: 184.00 Matches: 66
Percent Similarity: 38.89% Conservative: 11
Best Local Similarity: 33.33% Mismatches: 58
Query Match: 13.16% Indels: 63
DB: Gaps: 10

US-10-023-182-1 (1-752) x CA21_CANFA (1-1366)
QY 412 CTGCGACCGGAGACGGTGGCATCTGG-----CCAGGCTCTGCGGCGCAGCTCT 359
DB 630 ProGlyThrAlaGlyAlaSerGly-ProGlyGlyLeuProGlyLysAlaAlaGI 649
QY 358 GCTTCATGGGTGTGCGAAGGCGAGGAGGTAGACTCAG----- 315
DB 649 yLlePro-GlyGlyLysGlyGluLysGlyGluThrGlyLeuArgGlyLysGI 669
QY 314 -----CAGCGGCTCTCGGCGCCCGCGCATCTG 281
DB 669 roGlyArgAspGlyAlaArgGlyAlaProGlyAlaMetGlyAlaProGlyProAlaGlyA 689
QY 280 CAGCATCATTCAGCCCTGAGCGCGCGC-----CCATGCGGAGCCCGC 236
DB 689 lathrglyAspArg-GlyGluAlaGlyProAlaGlyProAlaGlyProAlaGlyProArg 708
QY 235 GGGGCGCT-----CTCCCGGCGCGGAGGCGCTTGTCTGCCCT 197
DB 709 GlyThrProGlyLysArgGlyLysAlaGlyProAlaGlyProAsnGlyPheAlaGlyPro 728
QY 196 GCG----- 194
DB 729 AlaGlyAlaAlaGlyLysProGlyAlaLysGlyLysArgGlyThrLysGlyProLysGI 748
QY 193 -----CCCGGAGACCTTGCAGCCCGCGCGCATCTCTTCTGCGCGCGAGCA 143
DB 749 GluAsnGlyProValGlyProThrGlyProIleGlySerAlaGlyProSerGlyProAsn 768
QY 142 TTGGCGCGCGCGCGCATCAGGAGT-----CAGGGGCGCTCTGGG----- 104
DB 769 GlyProProGlyProAlaGlySerArgGlyAspGlyLysProProGlyAlaThrGlyPhe 788
QY 103 CCATGACATGCGCCGCTGCAACCCCTGTCGCGCGCTTGCAGTGCATGCTCCGAG 44
DB 789 ProGlyAlaAlaGlyArgThrGlyPro--ProGlyProSerGlyLysThrGlyProPro 807
QY 43 CCTCTGCGCGCGCTCTGAG--AGAGAGGT--CAGGGCGCGAGAGCA 2
DB 808 GlyProProGlyAlaAlaGlyLysGlyLysGlyLysGlyLysGlyLysGlyLysGlyLys 823

RESULT 11
Fml4_MOUSE

```


RT "The human type I collagen mutation database.";
RT Nucleic Acids Res. 25:181-187(1997).
RN
RP VARIANT OI-II CYS-1166.
RX MEDLINE=86287390; PubMed=3016737.
RA Conn D.H., Byers P.H., Steinmann B., Gellinas R.E.;
RT "lethal osteogenesis imperfecta resulting from a single nucleotide
RT change in one human pro alpha 1(I) collagen allele.";
RT Proc. Natl. Acad. Sci. U.S.A. 83:6045-6047(1986).
RN
RN [15]
RP VARIANT OI-II ARG-569.
RX MEDLINE=87222295; PubMed=3108247.
RA Bateman J.F., Chan D., Walkers I.D., Rogers J.G., Cole W.G.;
RT "lethal perinatal osteogenesis imperfecta due to the substitution of
RT asparagine for glycine at residue 391 of the alpha 1(I) chain of type I
RT collagen.";
RL J. Biol. Chem. 262:7021-7027(1987).
RN
RN [16]
RP VARIANT OI-II CYS-926.
RX MEDLINE=88033031; PubMed=3667599;
RA Vogel B.E., Minor R.R., Freund M., Prockop D.J.;
RT "A point mutation in a type I procollagen gene converts glycine 748
RT of the alpha 1 chain to cysteine and destabilizes the triple helix in
RT a lethal variant of osteogenesis imperfecta.";
RL J. Biol. Chem. 262:14737-14744(1987).
RN
RN [17]
RP VARIANT OI-II ARG-842.
RX MEDLINE=88298828; PubMed=3403550;
RA Bateman J.F., Lennane S.R., Dahl H.H., Chan D., Cole W.G.;
RT "Substitution of asparagine for glycine 664 in the collagen alpha 1(I)
RT chain in lethal perinatal osteogenesis imperfecta. Demonstration of
RT the peptide defect by in vitro expression of the mutant cDNA.";
RL J. Biol. Chem. 263:11627-11633(1988).
RN
RN [18]
RP VARIANT OI CYS-1195.
RX MEDLINE=89218628; PubMed=3244312;
RA Labhard M.E., Witz M.K., Pope F.M., Nicholas A.C., Hollister D.W.;
RT "A cysteine for glycine substitution at position 1017 in an alpha
RT 1(I) chain of type I collagen in a patient with mild dominantly
RT inherited osteogenesis imperfecta.";
RL Mol. Biol. Med. 5:197-207(1988).
RN
RN [19]
RP VARIANT OI-II VAL-434.
RX MEDLINE=89255493; PubMed=2470760;
RA Patterson E., Smiley E., Bonadio J.;
RT "RNA sequence analysis of a perinatal lethal osteogenesis imperfecta
RT mutation.";
RL J. Biol. Chem. 264:10083-10087(1989).
RN
RN [20]
RP VARIANT OI-IV SER-1010.
RX MEDLINE=89308591; PubMed=2745420;
RA Martini J.C., Grange D.K., Gottesman G.S., Lewis M.B., Keoplin D.A.;
RT "Osteogenesis imperfecta type IV. Detection of a point mutation in
RT one alpha 1(I) collagen allele (COL1A1) by RNA/RNA hybrid analysis.";
RL J. Biol. Chem. 264:11893-11900(1989).
RN
RN [21]
RP VARIANTS OI-II ALA-1106; VAL-1151; ARG-1154 AND VAL-1184.
RX MEDLINE=89380165; PubMed=2777764;
RA Lennane S.R., Dahl H.-H.M., Cole W.G., Bateman J.F.;
RT "Characterization of point mutations in the collagen COL1A1 and
RT COL1A2 genes causing lethal perinatal osteogenesis imperfecta.";
RL J. Biol. Chem. 264:15809-15812(1989).
RN
RN [22]
RP VARIANT OI SER-1022.
RX MEDLINE=90062068; PubMed=2511192;
RA Pack M., Constantino C.D., Xiala K., Nielsen K.B., Prockop D.J.;
RT "Substitution of serine for alpha 1(I)-glycine 844 in a severe
RT variant of osteogenesis imperfecta minimally destabilizes the triple
RT helix of type I procollagen. The effects of glycine substitutions on
RT thermal stability are either position of amino acid specific.";
RL J. Biol. Chem. 264:19694-19699(1989).
RN
RN [23]
RP VARIANT OI-II CYS-1082.

RX MEDLINE=89109573; PubMed=2913053; Prockop D.J.;
 RA Constantinou C.D., Nielsen K.B.,
 RT "A lethal variant of osteogenesis imperfecta has a single base
 RT mutation that substitutes cysteine for glycine 904 of the alpha 1(I)
 RT chain of type I procollagen. The asymptomatic mother has an
 RT unidentified mutation producing an overmodified and unstable type I
 RT procollagen."
 RL J. Clin. Invest. 83:574-584(1989).
 RP [24]
 RP VARIANTS OF CYS-272; CYS-704 AND CYS-896.
 RX MEDLINE=90009313; PubMed=27824057.
 RA Steerman B.J., Eyre D., Charbonneau H., Harrylock M., Wells M.A.,
 RA Weiss L., Graham U.M., Byers P.H.:
 RT "Osteogenesis imperfecta: the position of substitution for glycine by
 RT cysteine in the triple helical domain of the pro alpha 1(I) chains of
 RT type I collagen determines the clinical phenotype."
 RL J. Clin. Invest. 84:1206-1214(1989).
 RP [25]
 RP VARIANT OF-II CYS-422.

Alignment Scores:		Pred. No.:	0.000152	Length:	1464
Score:	185.50	Matches:	87		
Percent Similarity:	36.86%	Conservative:	21		
Best Local Similarity:	29.69%	Mismatches:	52		
Query Match:	13.27%	Indels:	93		
DB:	1	Gaps:	13		

US-10-023-182-1 (1-752) x CAL1_HUMAN (1-1464)

QY	683	TCAGGGCCCCACAAATGAATGGCCACTGCTGTGGACCATTCCTTAGGGGAGAGGCAT	62	
DB	649	AlaGlyProProGlyValIuLa-----GlyLysProGly	655	
QY	623	GACCTCGAAGGGGCGCCAGCGCTTGAGCCTTCGCCCTCGAGGAGGAGGCTAGCCAA	566	
DB	660	GLUGInGlyValProGlyLysPleugly--AlaProGlyProSerGlyValAlaGlyGlu	678	
QY	563	AAA-----CACGGGAGAAAGACACTGGCTGATCCACATCAACAGGGAAGACTCTGGAG	510	
DB	679	ArgGlyPheProGlyValIuGlyValGInGlyPro-----ProGlyProAlaGlyPro	696	
QY	509	ACAGAGCTGATGAGAGCTGACAGTTGGCGGTGCTCCAGACAGTCACTCGATATGTCAG	450	
DB	697	ArgGlyValAlaSmGly--AlaProGlyValAsnAspGlyValAlaGlyValAspAlaGly	712	
QY	449	TATGTTCCGGACACAGTCACTCCTTCAAGAGCACCCCTGGACAGGAGAGCGGTGGGC	390	
DB	714	-----AlaProGlyValaProGlySerGInGlyA	722	
QY	389	ATCCTCGAG-----CCAGGCTCTCGGCGGCCACAGCTCGTTCATGAGGTGTGCG	342	
DB	723	laProGlyLysGInGlyMetProGlyGluArgGlyValAlaGlyLysProGlyProLys	742	
QY	341	GAAAGGCATGGCGAGGTGAAGTCAACAGCGCG-----	302	
DB	743	GlyAspArgGlyAspAlaGlyProLysGlyValaAspGlySerProGlyLysAspGlyVal	762	
QY	308	-----GCTCCGCGGCCCCCTGGGCCCGCATCGACAGATTCATCAGGCC	266	
DB	763	ArgGlyLysGInGlyProLysGlyProProGlyProAlaGlyValaProGlyLysPheGly	782	
QY	263	TGAAGCCGCGCGCG-----CCATGCGAGCCCGCGGGGCGCT-----	222	
DB	782	GlyLysGlyProSerGlyProAlaGlyProThrGlyValaArgGlyValaProGlyLysAsp	802	
QY	226	-----CTCCGCGCGCGGAGCCCTTCGCTCGCCCTCGCGCCCGCGGAGCTCT	180	
DB	802	GlyGlyLysProGlyProProGlyProAlaGlyLysPheAlaGlyPro-----ProGlyValaAspGly	822	
QY	179	GCGGCGCGTGGCACCCGCGCTTCCTGGGCGG-----CCAGCATTTGCCCC	138	
DB	821	GlyLysProGlyValaGlyLysGlyLysProGlyValaAspAlaGlyLysGlyValaAspAlaGlyPro	841	

MGD, MGI:88467, Col1a1.
 DR InterPro: IPR00087, Collagen.
 DR InterPro: IPR00885, Fib_collagen_C.
 DR InterPro: IPR01007, vwf_C.
 DR Pfam: PF01410, COLFI, 1.
 DR Pfam: PF01391, Collagen: 18.
 DR ProDom: PD000007, C1g_helix, 1.
 DR ProDom: PD002878, Fib_collagen_C, 1.
 DR SMART: SM00038, COLFI, 1.
 DR SMART: SM00214, VWC, 1.
 DR PROSITE: PS01208, VWC, 1; 1.
 DR PROSITE: PS50184, VWC, 2; 1.
 KM Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 KM Glycoprotein; Collagen; Signal.
 FT SIGNAL 1 22
 FT PROPEP 23 151 AMINO-TERMINAL PROPEPTIDE.
 FT CHAIN 152 1207 COLLAGEN ALPHA 1(1) CHAIN.
 FT PROPEP 1208 1453 CARBOXYL-TERMINAL PROPEPTIDE.
 FT DOMAIN 29 87 VWC.
 FT DOMAIN 152 167 NONHELICAL REGION (N-TERMINAL).
 FT DOMAIN 168 181 TRIPLE-HELICAL REGION.
 FT DOMAIN 1182 1207 NONHELICAL REGION (C-TERMINAL).
 FT CARBOHYD 56 56 N-LINKED (GLUCNA. . .) (POTENTIAL).
 FT CARBOHYD 1354 1354 N-LINKED (GLUCNA. . .) (POTENTIAL).
 FT SITE 734 736 CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 1082 1084 CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 1450 1450 A -> V (IN REF. 5).
 SQ SEQUENCE 1453 AA; 137944 MW; 3B802E55D81808 CRC64;
 Alignment Scores:
 Pred. No.: 0.000107 Length: 1453
 Score: 188.00 Matches: 66
 Percent Similarity: 37.31% Conservative: 9
 Best Local Similarity: 32.84% Mismatches: 62
 Query Match: 13.45% Indels: 64
 Gaps: 8
 DB: 1
 US-10-023-182-1 (1-752) x CA11_MOUSE (1-1453)
 QY 417 GCACCCCTGGACGGAGCGGTGGGCACTCTCTGG-----CGAGCTCTCTCG 370
 DB 703 A1APROGLYALAProglyserGlnGlyLAProGlyLeuGlnGlyMeProGlyGluArg 722
 QY 369 GGGCAGCTCTGCTTCATCGAGGTGTCGCCAAGAGCATGGAGTGAAGTCAAGCAGAGC 310
 DB 723 G1yA1aA1aG1yLeuPro-G1yProLyse1yAspArg1yAspAlaG1yProLyse1yAl 742
 QY 309 G-----GCTCTCCGCGCCCTCG 292
 DB 742 AAspGlyserProGlyLyAspGlyLA1aArg1yLeuThrg1yPro1leG1yProProG1 762
 QY 291 CCCCCGATCTGCAGCATTCATTCAGCCCGAAGCCGCGCG-----CGAT 247
 DB 762 yProAlaG1yLAProG1yAspLyS-G1yGluA1aG1yProserG1yProProG1yProT 782
 QY 246 GCGGACCCCGCGGGCGCGCT-----CCTCCCGGCGCCGAGGCGCC 208
 DB 782 hGc1yLA1aArg1yLA1aProG1yAspArg1yGluA1aG1yProProG1yProAlaG1yP 802
 QY 207 TTGCTGCCCCCT----- 197
 DB 802 hE1aG1yProProG1yLA1aAspG1yGlnProG1yLA1aLyse1yGluProG1yAspThrg 822
 QY 196 -----GCGCCCGGAGCACTCTGCGCGCGCGTGGACCCGCGCTCTCT- 155
 DB 822 1yVal1yse1yAspAlaG1yProProG1yProAlaG1yProAlaG1yProProG1yProI 842
 QY 154 -----GGCGCGCCGCACTGCGCCCGCATCAGAAATGCCAGGGCTCTCTG 106
 DB 842 1eG1yAsnValG1yLA1aProG1y---ProLyse1yProArg1yLA1aA1aG1yProProG 861
 QY 105 GGCATCAGCATGCGCC-----GTCAACCCCGCTGAGCCCGGCGCTTCGG 61

DB 861 1yLA1aThrg1yheProG1yLA1aA1aG1yArgValG1yProProG1yProserG1yAsn 881
 QY 60 CCGCATGCTCCGAGCGCTCTGCGCGCTCTGCAGAGGAAGTGCAGGCGCCACAGAGA 2
 DB 881 1aG1yProProG1yProProG1yProValG1yLyse1yG1yG1yLyse1yProArg1y 900
 RESULT 7
 ID HYDL_STRCO STANDARD; PRT; 627 AA.
 AC P42534, Q95217.
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative polyketide hydroxylase (BC 1.14.13.-) (WHE ORF VII).
 GN SC05321 OR SC69.12C.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wierczek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.,
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RL coelicolor A3(2)."
 RN Nature 417:141-147(2002).
 RP SEQUENCE OF 1-255 FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=94075247; PubMed=8253693;
 RA Bianco G., Ferrel A., Brian P., Mendez C., Chater K.F., Salas J.A.;
 RT "A hydroxylase-like gene product contributes to synthesis of a
 RL polyketide spore pigment in Streptomyces halstedii.";
 RJ J. Bacteriol. 175:8043-8048 (1993).
 CC -1- FUNCTION: INVOLVED IN DEVELOPMENTALLY REGULATED SYNTHESIS OF A
 CC COMPOUND BIOSYNTHEMICALLY RELATED TO POLYKETIDE ANTIBIOTICS,
 CC WHICH IS ESSENTIAL FOR SPORE COLOR IN STREPTOCOCCUS COELICOLOR.
 CC -1- CORPACTORY: FAD (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PHEA/TPDB FAMILY OF FAD MONOOXYGENASES.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DB EMBL; AL39123; CAB45603.1; -
 DB EMBL; X74213; CA52289.1; -
 DB PIR; T35608.
 DR InterPro: IPR000087, Collagen.
 DR InterPro: IPR000733, Flav_monooxygenase.
 DR InterPro: IPR003938, Moxy_FAD_binding.
 DR InterPro: IPR003042, Rng_mnoxygenase.
 DR Pfam: PF01494; FAD_binding_3; 1.
 DR Pfam: PF01360; Monooxygenase; 1.
 DR PRINTS; PR00420; RINGMOXGNASE.
 KM Oxidoreductase; Flavoprotein; FAD: Complete proteome.
 FT NP_BIND 22 51 FAD (POTENTIAL).
 FT NP_BIND 309 319 FAD (POTENTIAL).
 FT NP_BIND 60 60 R -> A (IN REF. 2).
 FT CONFLICT 145 145 L -> LH (IN REF. 2).
 FT CONFLICT 234 234 C -> S (IN REF. 2).

US-10-023-182-1 (1-752) x GPI_CHLRE (1-555)

CALL_MOUSE	STANDARD	PRT	1453 AA.
AC	P11087; Q60635;		
DT	01-JUL-1989 (Rel. 11, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Collagen alpha 1(I) chain precursor.		

GN COL1A1 OR COL1A1..
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; E
 OC Mammalia; Eutheria; Rodentia; Scurionath; Muridae;
 CX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N;
 RX MEDLINE=96033240; Pubmed=9535610;
 RA Li S.W., Khillan J., Prockop D.J.;
 RT The complete cDNA coding sequence for the mouse pro alpha 1(I) chain
 RT of type I procollagen.";
 RL Matrix Biol. 14:593-595 (1995).
 RN [2]
 RP SEQUENCE OF 518-1128 FROM N.A.
 RX MEDLINE=86137403; Pubmed=3841523;
 RA French B.T., Lee W.-H., Maul G.G.;
 RT "Nucleotide sequence of a cDNA clone for mouse pro alpha 1(I)
 RT collagen protein.";
 RL Gene 39:311-312 (1985).
 RN [3]
 RP SEQUENCE OF 735-1130 FROM N.A.
 RX MEDLINE=83141374; Pubmed=6298597;
 RA Monson U.M., Friedman J., McCarthy B.J.;
 RT "DNA sequence analysis of a mouse pro alpha 1 (I) procollagen gene:
 RT evidence for a mouse B1 element within the gene.";
 RL Mol. Cell. Biol. 2:1362-1371 (1982).
 RN [4]
 RP SEQUENCE OF 735-878 AND 1005-1058 FROM N.A.
 RX MEDLINE=83157109; Pubmed=6219867;
 RA Monson U.M., McCarthy B.J.;
 RT "Identification of a Balb/c mouse pro alpha 1(I) procollagen gene:
 RT evidence for insertions or deletions in gene coding sequences.";
 RL DNA 1:59-69 (1981).
 RN [5]
 RP SEQUENCE OF 1442-1453 FROM N.A.
 RX MEDLINE=86124276; Pubmed=3340560;
 RA Wooslemer K., Haverz K.;
 RT "Two mRNAs of mouse pro alpha 1(I) collagen gene differ in the size
 RT of the 3'-untranslated region.";
 RL Nucleic Acids Res. 16:1773-773 (1988).
 CC -1- FUNCTION. TYPE I COLLAGEN IS A MEMBER OF GROUP I COLLAGEN
 CC (FIBRILLAR FORMING COLLAGEN).
 CC -1- SUBUNIT. TRIMERS OF ONE ALPHA 2(I) AND TWO ALPHA 1(I) CHAINS.
 CC -1- TISSUE SPECIFICITY. FORMS THE FIBRILS OF TENDON, LIGAMENTS AND
 CC BONES. IN BONES THE FIBRILS ARE MINERALIZED WITH CALCIUM
 CC HYDROXYAPATITE.
 CC -1- PMW: Prolines at the third position of the tripeptide repeating
 CC unit (G-X-Y) are hydroxylated in some or all of the chains.
 CC -1- SIMILARITY: Contains 1 WPC domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U08020; AAA88912.1; -
 CC EMBL; X15896; CA333904.1; -
 CC EMBL; M14423; AAA37333.1; -
 CC EMBL; M17491; AAA37334.1; -
 CC EMBL; X06753; CA29927.1; -
 CC EMBL; X03036; AAA37332.1; -
 CC EMBL; K03029; AAA37332.1; JOINED.
 CC EMBL; K03030; AAA37332.1; JOINED.
 CC EMBL; K03031; AAA37332.1; JOINED.
 CC EMBL; K03032; AAA37332.1; JOINED.
 CC EMBL; K03033; AAA37332.1; JOINED.
 CC EMBL; K03034; AAA37332.1; JOINED.
 CC EMBL; K03035; AAA37332.1; JOINED.
 CC PIR; S57243; S21626.

215 ProglyAlnProGlyProProGlyGlnAlaGlyProGInGlyProProcGly 231

```
RESULT 4
WASP MOUSE
ID       WASP_MOUSE      STANDARD;          PRT;           520 AA.
AC       P70315;
DT       01-NOV-1997 (Rel. 35, Created)
DD       01-NOV-1997 (Rel. 35, Last sequence update)
DE       28-FEB-2003 (Rel. 41, Last annotation update)
DEFINITION   Wiskott-Aldrich syndrome protein homolog (WASP).
SOURCE        MMS OR WASP.
ORGANISM      Mus musculus (Mouse).
OC            Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;
OC            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX           _Taxid=10090;
RN           [1]
SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=P6115600; PubMed=8666397;
RA Dery J.M., Wiemann P., Blair P., Yang Y., Kerns J.A.,
RA Lemahieu V., Godfrey V.L., Wilkinson J.E., Francke U.;
RT "The mouse homolog of the Wiskott-Aldrich syndrome protein (WASP)
RT gene is highly conserved and maps near the scurfy (sf) mutation on
RT the X chromosome." ;
RL Genomics 29:471-477(1995).
CC -! FUNCTION: POSSIBLE REGULATOR OF LYMPHOCYTE AND PLATELET FUNCTION.
CC MAY BE INVOLVED IN SIGNALING PATHWAYS WITH CYTOSKELETTAL FUNCTION
CC (BY SIMILARITY).
CC -! DOMAIN: THE WH1 (WASP HOMOLOGY 1) DOMAIN MAY BIND A PRO-RICH
CC LIGAND.
CC -! SIMILARITY: Contains 1 CRIB domain.
CC -! SIMILARITY: Contains 1 WH1 domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isd-sib.ch/announce/isb-sib.ch).
or send an email to license@isb-sib.ch).
-----
DR EMBL; U54788; AAC52556.1; -.
DR MGD; MG1:105059; Was.
DR InterPro; IPR0000697; EVHL.
DR InterPro; IPR000095; PABox/RhoBindng.
DR InterPro; IPR001960; WH1.
DR InterPro; IPR001324; WH2.
DR Pfam; PF00786; PBD; 1.
DR Pfam; PF00568; WH1; 1.
DR Pfam; PF02205; WH2; 1.
DR SMART; SMO0265; PBD; 1.
DR SMART; SMO0461; WH1; 1.
DR SMART; SMO0246; WH2; 1.
DR PROSITE; PS0108; CRIB; 1.
DR Repeat; Phosphorylation.
KW FMK
FT DOMAIN             41              WH1.
FT REPEAT             240            253    CRIB.
FT REPEAT             354            363    GRSGLPXPX MOTIF 1.
FT REPEAT             393            402    GRSGLPXPX MOTIF 2.
FT DOMAIN             162            167    POLY-PRO.
FT DOMAIN             314            321    POLY-PRO.
FT DOMAIN             324            321    POLY-PRO.
FT DOMAIN             368            373    POLY-GLY.
FT DOMAIN             376            379    POLY-PRO.
FT DOMAIN             384            390    POLY-PRO.
FT DOMAIN             397            403    POLY-PRO.
FT DOMAIN             408            424    POLY-PRO.
FT DOMAIN             503            520    ASP/GLU-RICH (ACIDIC).
FT MOD RES             293            293    PHOSPHORYLATION (BY SIMILARITY).
FT SEQUENCE           520 AA;         54191 MW;          9C22373GC59FOC8A CRC64;
```

Alignment Scores:

[illegible]

```

QY      174 GCGCGCAGAGGTCCTCCGGGGGCCCAAGGGCAGCAGGGCTTGGGGCCGGAGAGAGTGCCC    233
Db      41  GLVGLYARGGLYProArgGLYAlaGLYAlaAlaAgAlaSerGLYProArgGLYGLYAla    60
QY      234 CCGCGGGAGTCCCGCATGGCCGCCTTCAGGGGTGAATGANTCTCAGATGCGGGGCC    293
Db      61  ProArgGLYProHISetLYLAlaAlaSerAlaGlnApeGLYArgCySProCYSGLYAla    80
QY      294 AGGGGGCCGGAGAGCCGCTTGTTAAGTTCACTCCGCATGCCCTTTCGGACACCAG    353
Db      81  ArgArgPAsperBerArgLeuLeuGluLeuHISleIleHmETProPheSerSerPromer    100
QY      354 GAAGCAGAGCTGGCCCCCAGAGAGCTTGGCCCCAGATGSCCCACCGCTTCCCGTCCAAGG    413
Db      101 GLVALAGLUenValArgArgILeUSeSerArgApeAlaAlaProLeuProArgProGLY    120
QY      414 GTGCTTCTGAAGAAGTTCACTGTGTTCGGGCAACAATCAGATACATACGACTGACGTCTGCA    473
Db      121 AlValLeuYeasPhenThrValSerGlyLeuLeuLeuPheMetSerValTPAspGln    140
QY      474 GACCACCGCCCACTGCAGCTTCATCAGCTCTGTCTCCAGCAGCTTTCCTGTGTGATG    513
Db      141 AspArgGLUGLY---:::-----AlaGlyArgetArgValAlaGLY    152
QY      534 TGGATCACCGCAGTGCTTTCGCGCCGTGTTTTGGCTGACGCTCCCTCAGGGCAGAGS---    590
Db      153 TRP-----GLyeuGLysSerAlaSerProGLUGLYGlnYSAla    165
QY      591 -----CGCTAAAGCCACGCTTGAGCGCCCTTCCTAG    620
Db      166 ArgAspLeuArgThzProLySHISyValSerGlnGlnArgProGLYThnProGLY---    184
QY      621 GTCATGCTCTCCCTCCCTAGG    641
Db      185 ---ProProProProGLUGLY    190

RESULT 3
CC34_CABEL ID CC34_CABEL STANDARD; PRT; 299 AA.
AC P346B7: Q20087;
DT 01-FEB-1994 (Rel. 28, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cuticle collagen 34 (Abnormal ray morphology protein 4).
GN COL-34 OR RAM-4 OR F36A4.10.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;
OC Rhabditiidae; Peloderiinae; Caenorhabditis.
ON NCBI_TaxID=6239;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=93013043; PubMed=1398138;
RA Bird D.M.;
RT "Sequence comparison of the Caenorhabditis elegans dpy-13 and col-34 genes, and their deduced collagen products.";
RL Gene 120:261-266(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX Miller N., Bradshaw H.;
RL Submitted (APR-1996) to the EMBL/genbank/DBJ databases.
CC -! FUNCTION: NEMATODE CUTICLES ARE COMPOSED LARGELY OF COLLAGEN-LIKE PROTEIN. THE CUTICLE FUNCTIONS BOTH AS AN EXOSKELETON AND AS A BARRIER TO PROTECT THE WORM FROM ITS ENVIRONMENT.
CC -! SUBUNIT: COLLAGEN POLYPEPTIDE CHAINS ARE COMPLEXED WITHIN THE CUTICLE BY DISULFIDE BONDS AND OTHER TYPES OF COVALENT CROSS-
```

```
CC LINKS:-----
CC -1- SIMILARITY: BELONGS TO THE CUTICULAR COLLAGEN FAMILY.-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@sib-sib.ch).-----
CC DR EMBL; M80650; AAA27985.1; .
CC DR EMBL; AF410845; AAL76233.1; .
CC DR EMBL; U53333; AAA96155.1; .
CC DR PIR; JCI448; JCI448.
CC DR PIR; T29956; T29956.
CC DR Wormpep; F16A4.10; C807185.
CC DR GO; GO:0040002; Peptide Biosynthesis (sensu Nematoda); NAS.
CC DR Interpro; IPR002486; Col_cuticle_N.
CC DR Interpro; IPR000087; Collagen.
CC DR Pfam; PF01484; Col_cuticle_N; 1.
CC DR Pfam; PF01391; Collagen_2.
CC KW Cuticle; Connective tissue; Repeat; Multigene family; Collagen.
CC FT DOMAIN 216 278 TRIPLE-HELICAL REGION.
CC FT CONFLICT 145 145 C -> K (IN REF. 1).
CC FT CONFLICT 189 189 P -> W (IN REF. 1).
CC FT CONFLICT 199 210 GFPGGVVFAQS -> EHGSQECPRG (IN REF. 1).
CC SQ SEQUENCE 299 AA; 29336 MW; 2157808085C7212 CRC64;
US-10-023-182-1 (1-752) x CC34_CABRL (1-299)
OY 492 GCGAGCTGGGGGAGGTCTGCAGACAGTCAGATGATCATGATTGCCGACAAG 433
DB 84 ALALVALTHGLYGLYALASINSERLYSERYSLSUERYCYSLSEU----- 101
OY 432 TGAAGCTCTTGCAAAAGCACCCCTGGACAGGAGAAGCGGTGGGACACTCTGGGCCAGGCTCC 373
DB 102 -----ProglyPro 104
OY 372 TGCGGCGACAGCTCTGCTTCATGAGGTGTGCGAAGAGCATGGCGAGGTAAATCAAGCA 313
DB 105 ProglyProAlaglyThrPro-----GlySerProGlyArgProGlyLysProGlyAla 122
OY 312 GCGGCTCTCGGC-----CCCTGCCCCGATCTGACATGCATTCATCAAGCC 257
DB 123 ProglyLeuProglyAsnProglyValArgProPro--GlnInProCysGluProIleThr 141
OY 256 GCGCGCCGACATGGAGAACCCGCGGGCGGCTCTCCCGGACCAGGCGCTTGCTGCCCCT 197
DB 142 ProProProCylProProCysProProGlnLysProProGlyPro-----Pro 156
OY 196 GCGCCCCGGGGAACCTCTGCGCGCGGTGGACCCGCTCTCTGGGCGCGACAGATTGCC 137
DB 157 GlyProProgly-----ProProglyAspSerGlyGluProGlySerProglyLeuPro 174
OY 136 -----CCTGGGCGATCAAGAAATGCGAGAGGCTCTCGG6--- 104
DB 175 GlyGlnAspAlaAlaProglyGluInProglyProLysGlyProProglyProProglyAla 194
OY 103 CGATCAAGCA-----TCGCGCGTGAACCCCTGTGCCCGCGGCTTGCGCTGCAATGGCT 50
DB 195 ProGlyAlaProglyThrProglyGluInProglyValProAlaGlnSerGluInProLeuIle 214
OY 49 CGC----GAGCCTCTGCGCGCTCTCAAGAGAAAGTGTAGAGGACCACAGAGA 2
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 10, 2004, 17:29:01 ; Search time 14.5 Seconds
(without alignments)
4877.804 Million cell updates/sec

Title: US-10-023-182-1

Perfect score: 1403
Sequence: 1 ATCTCTGCTGGGCTGACCT.....TAAACTGACCTACGAAAA 752

Scoring table:

BIOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delcp 6.0 , Delcxt 7.0

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 255726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ n2p model -DEV=xlp
-O=/cgm2.1/USPRO_spool_p/US10023182/runat_10022004_171428_3863/app_query.fasta.1.903
-DB=SwissProt 41 -QFMT=fastan -SUFFIX=esp -MINMATCH=0.1 -DOOPL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLGN=200 -THR_SCORE=spct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USPR=US10023182 @CGN 1.1.16 @runat 10022004_171428_3863 -NCPU=6 -ICPU=3
-NO MAP -LARGEOUTRY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGIOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	959	68.4	180	1 CTG1_HUMAN	P78358 homo sapien
2	597	42.6	210	1 CTG2_HUMAN	O75638 homo sapien
3	196	14.0	299	1 CC34_CABEL	P34667 caenorhabdi
4	192.5	13.8	520	1 WASP_MOUSE	P70315 mus musculu
5	192	13.7	555	1 GPI_CHURE	O9f966 chlamydomon
6	188	13.4	1453	1 CAL1_MOUSE	P11087 mus musculu
7	186.5	13.3	627	1 HYDL_STRCO	P42534 streptomyce
8	185.5	13.3	1464	1 CAL1_HUMAN	P02452 homo sapien
9	184	13.2	779	1 CAL1_BOVIN	P02453 bos taurus
10	184	13.2	1366	1 CA21_CANPA	O46382 canis famli
11	183	13.1	1206	1 FMN1_MOUSE	O05859 mus musculu
12	183	13.1	1468	1 FMN1_MOUSE	O05860 mus musculu
13	182	13.0	1364	1 CA21_BOVIN	P02465 bos taurus
14	181	12.9	296	1 CC01_CABEL	P08124 caenorhabdi
15	180.5	12.9	1460	1 CAL1_CANPA	O9x517 canis famli
16	179	12.8	1049	1 CAL1_BOVIN	P04358 bos taurus
17	178.5	12.8	301	1 CC02_CABEL	P17656 caenorhabdi
18	177	12.7	1372	1 CA21_RAT	P02456 rattus norv

C 19	176.5	12.6	261	1 PRP2_MOUSE	P05142 mus musculu
C 20	176.5	12.6	266	1 YXWK_CABEL	O21184 caenorhabdi
C 21	176.5	12.6	296	1 PRP3_MOUSE	P05143 mus musculu
C 22	175.5	12.6	1790	1 SEPA_EMENI	P78621 emericella
C 23	174.5	12.4	641	1 ENR1_BHV	P03121 epstein-bar
C 24	173.5	12.4	1366	1 CA21_HUMAN	P08123 homo sapien
C 25	173.5	12.4	1049	1 CA13_BOVIN	P04258 bos taurus
C 26	171.5	12.3	1690	1 CA44_HUMAN	P53420 homo sapien
C 27	171	12.2	675	1 CA39_CHICK	P32017 gallus galli
C 28	170.5	12.2	1464	1 CA13_MOUSE	P08121 mus musculu
C 29	170	12.2	534	1 APG_ABAT	P40602 arabidopsis
C 30	170	12.2	1453	1 CA11_CHICK	P02457 gallus galli
C 31	168.5	12.0	1356	1 CA21_ONCMY	O91484 oncornychu
C 32	168	12.0	505	1 WMSL_BOVIN	O95107 bos taurus
C 33	168	12.0	627	1 SPD2_NEPCL	P48804 nephila cia
C 34	167.5	11.9	779	1 CA11_BOVIN	P02453 bos taurus
C 35	167.5	11.9	1464	1 CA11_HUMAN	P02452 homo sapien
C 36	167	11.9	1466	1 CA13_HUMAN	P02461 homo sapien
C 37	166.5	11.9	684	1 CA39_HUMAN	O14050 homo sapien
C 38	166.5	11.9	1460	1 CA11_CANPA	O9x517 canis famli
C 39	166	11.9	1083	1 T2D3_HUMAN	O00268 homo sapien
C 40	166	11.8	1466	1 CA13_HUMAN	P02461 homo sapien
C 41	166	11.8	3119	1 CA1C_MOUSE	O60847 mus musculu
C 42	165.5	11.8	485	1 SSGP_VOLICA	P21997 volvox cart
C 43	165	11.8	302	1 CCDC_CABEL	P17657 caenorhabdi
C 44	165	11.8	1372	1 CA21_MOUSE	O01149 mus musculu
C 45	164	11.7	1336	1 W146_HUMAN	O9c058 homo sapien

ALIGNMENTS

RESULT 1	CTG1_HUMAN	STANDARD;	PRT;	180 AA.
ID	CTG1_HUMAN			
AC	P78358;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Cancer/testis antigen 1 (Autoimmunogenic cancer/testis antigen NY-ESO-1).			
GN	CTAG1 OR CTAG.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97203161; PubMed=9050879;			
RA	Chen Y.-T., Scanlan M.U., Sahin U., Tuercel O., Gure A.O., Tsang S.,			
RA	Williamson B., Stockert E., Pfeundschun M., Old U.J.;			
RT	"A testicular antigen aberrantly expressed in human cancers detected			
RT	by autologous antibody screening."			
RL	Proc. Natl. Acad. Sci. U.S.A. 94:1914-1918(1997).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Melanoma;			
RX	MEDLINE=98289682; PubMed=9626360;			
RA	Lethie B., Lucas S., Michaux L., de Smet C., Godelaine D., Serrano A.,			
RA	de Plaen E., Boon T.;			
RT	"LAG3-1, a new gene with tumor specificity."			
RT	Int. J. Cancer 76:903-908(1998).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98430682; PubMed=9759882;			
RA	Wang R.-F., Johnston S.U., Zeng G., Topalian S.L.,			
RA	Shankar-Gupte D., Rosenberg S.A.;			
RT	"A breast and melanoma-shared tumor antigen: T cell responses to			
RT	antigenic peptides translated from different open reading frames."			
RT	J. Immunol. 161:3596-3606(1998).			
CC	- TISSUE SPECIFICITY: EXPRESSED IN TESTIS AND OVARY AND IN A WIDE			
CC	VARIETY OF CANCERS. DETECTED IN TESTIS AND OVARY AND IN A WIDE			
CC	- SIMILARITY: BELONGS TO THE CTAG FAMILY.			

Fri Feb 13 14:37:37 2004

us-10-023-182-4.ral

Page 6

```
/ Sequence 8, Application US/09751798
/ Patent No. 6525177
/ GENERAL INFORMATION:
/ APPLICANT: Stockert, Elisabeth; Jager, Elke;
/ APPLICANT: Chen, Yao-tseng; Scanlan, Matthew;
/ APPLICANT: Knuth, Alexander; Old, Lloyd U.
/ TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer
/ TITLE OF INVENTION: Associated Proteins, Uses Thereof,
/ TITLE OF INVENTION: Truncated Forms of NY-ESO-1, and HLA
/ TITLE OF INVENTION: Binding Peptides Derived Therefrom
/ NUMBER OF SEQUENCES: 8
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Fulbright & Jaworski, L.L.P.
/ STREET: 666 Fifth Avenue
/ CITY: New York City
/ STATE: New York
/ COUNTRY: USA
/ ZIP: 10103
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
/ OPERATING SYSTEM: PC-DOS
/ SOFTWARE: WordPerfect
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/751,798
/ FILING DATE:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 09/062,422
/ FILING DATE:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/725,182
/ FILING DATE: October 3, 1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Hanson, No. 6525177man D.
/ REGISTRATION NUMBER: 30,946
/ REFERENCE/DOCKET NUMBER: LUD 5466.3
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 318-3168
/ TELEFAX: (212) 752-5958
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 180
/ TYPE: amino acid
/ TOPOLOGY: linear
/ US-09-751-798-8

Query Match 100.0%; Score 61; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.0053;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLMMITQCFI 11
| | | | | | | | | |
Db 157 SLMMITQCFI 167
```

Search completed: February 10, 2004, 17:20:31
Job time : 14.4138 secs

ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/791,495
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Van Amsterdam, John R.
 REGISTRATION NUMBER: 40,212
 REFERENCE/DOCKET NUMBER: L0461/7005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 180 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-791-495-7

Query Match 100.0%; Score 61; DB 2; Length 180;
 Best Local Similarity 100.0%; Pred. No. 0.0053;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLWMITQCF 11
 DB 157 SLWMITQCF 167

RESULT 13

US-08-791-495-9
 Sequence 9, Application US/08791495
 Patent No. 581519
 GENERAL INFORMATION:
 APPLICANT: Leuch, Bernard
 APPLICANT: Lucas, Sophie
 APPLICANT: De Smet, Charles
 APPLICANT: Godelaine, Daniele
 APPLICANT: Boon-Falleur, Thierry
 TITLE OF INVENTION: L1-1 TUMOR SPECIFIC GENES
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/791,495
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Van Amsterdam, John R.
 REGISTRATION NUMBER: 40,212
 REFERENCE/DOCKET NUMBER: L0461/7005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500

TELEFAX: 617-720-2441
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 180 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-791-495-9

Query Match 100.0%; Score 61; DB 2; Length 180;
 Best Local Similarity 100.0%; Pred. No. 0.0053;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLWMITQCF 11
 DB 157 SLWMITQCF 167

RESULT 14

US-08-937-263B-8
 Sequence 8, Application US/08937263B
 Patent No. 6274145
 GENERAL INFORMATION:
 APPLICANT: Chen, Yao-Tseng; Scanlan, Matthew;
 APPLICANT: Gure, Ali; Old, Lloyd J.; Jager, Elke;
 APPLICANT: Alexander, Knuth; Drijfhout, Jan W.
 TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE
 TITLE OF INVENTION: ENCODING CANCER ASSOCIATED ANTIGEN, THE ANTIGEN
 TITLE OF INVENTION: ITSELF, AND USES THEREOF
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fulbright & Jaworski, L.L.P.
 STREET: 666 Fifth Avenue
 CITY: New York City
 STATE: New York
 COUNTRY: USA
 ZIP: 10103
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
 COMPUTER: IBM
 OPERATING SYSTEM: PC-DOS
 SOFTWARE: WordPerfect
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/937,263B
 FILING DATE: September 15, 1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/725,182
 FILING DATE: October 3, 1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Sinn, Eric, Patent Agent
 REGISTRATION NUMBER: 40,177
 REFERENCE/DOCKET NUMBER: LUD 5466.1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 318-3000
 TELEFAX: (212) 752-5598
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 180
 TYPE: amino acid
 TOPOLOGY: linear

US-08-937-263B-8

Query Match 100.0%; Score 61; DB 3; Length 180;
 Best Local Similarity 100.0%; Pred. No. 0.0053;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLWMITQCF 11
 DB 157 SLWMITQCF 167

RESULT 15
 US-09-751-798-8

```
RESULT 9
US-09-751-798-4
; Sequence 4, Application US/09751798
; Patent No. 6525177
; GENERAL INFORMATION:
; APPLICANT: Stoeckert, Elisabeth; Jager, Elke;
; APPLICANT: Chen, Yao-tseng; Scanlan, Matthew;
; APPLICANT: Knuth, Alexander; Old, Lloyd J.
; TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer
; TITLE OF INVENTION: Associated Proteins, Uses Thereof,
; TITLE OF INVENTION: Truncated Forms of NY-ESO-1, and HLA
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski, L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/751,798
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/062,422
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/725,182
; FILING DATE: October 3, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6525177man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5466.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3168
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-09-751-798-4
Query Match 100.0%; Score 61; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLMMITQCF 11
DB 1 SLMMITQCF 11
RESULT 10
US-09-344-040C-130
; Sequence 130, Application US/09344040C
; Patent No. 6548064
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfreundschuh, Michael
; APPLICANT: Ramensee, Hans Georg
; APPLICANT: Stevanovic, Stefan
; TITLE OF INVENTION: Method for Determining Presence of Cancer in a Sample By Determin
; TITLE OF INVENTION: Expression of an SSX Gene, Peptides Derived From Said SSX Gene a
; FILE REFERENCE: LUD 5556.1
; CURRENT APPLICATION NUMBER: US/09/344,040C
; CURRENT FILING DATE: 1999-06-25
```

```
;; PRIOR APPLICATION NUMBER: US 09/105,839
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: US 08/851,130
;; PRIOR FILING DATE: 1997-05-05
;; NUMBER OF SEQ ID NOS: 132
;; SEQ ID NO 130
;; LENGTH: 11
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-344-040C-130
Query Match 100.0%; Score 61; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLMMITQCF 11
DB 1 SLMMITQCF 11
RESULT 11
US-09-574-749B-33
; Sequence 33, Application US/09574749B
; Patent No. 6548299
; GENERAL INFORMATION:
; APPLICANT: ROSENZWEIG, Michael
; APPLICANT: PYKEIT, Mark J.
; APPLICANT: SCADEN, David T.
; APPLICANT: POZNANSKY, Mark C.
; TITLE OF INVENTION: LYMPHOID TISSUE-SPECIFIC CELL PRODUCTION
; TITLE OF INVENTION: FROM HEMATOPOIETIC PROGENITOR CELLS IN THREE-DIMENSIONAL
; FILE REFERENCE: C1005/7012/KA/ERG
; CURRENT APPLICATION NUMBER: US/09/574,749B
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US 60/107,972
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: PCT/US99/26795
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: US 09/524,749
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Homo Sapiens source
US-09-574-749B-33
Query Match 100.0%; Score 61; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLMMITQCF 11
DB 1 SLMMITQCF 11
RESULT 12
US-08-791-495-7
; Sequence 7, Application US/08791495
; Patent No. 5811519
; GENERAL INFORMATION:
; APPLICANT: Leth, Bernard
; APPLICANT: Lucas, Sophie
; APPLICANT: De Smet, Charles
; APPLICANT: Godetaine, Daniele
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: L1-1 TUMOR SPECIFIC GENES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
```

Query Match 100.0%; Score 61; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLMMWITOCFL 11
DB 1 SLMMWITOCFL 11

RESULT 5
US-09-166-448-79
; Sequence 79, Application US/09166448
; Patent No. 6291430
; GENERAL INFORMATION:
; APPLICANT: Chaux, Pascal
; APPLICANT: Vantomme, Valérie
; APPLICANT: Stroobant, Vincent
; APPLICANT: Boon-Falleur, Thierry
; APPLICANT: Van der Bruggen, Pierre
; APPLICANT: Thielemans, Kris
; APPLICANT: Cortbals, Jurgen
; TITLE OF INVENTION: MAGE-3 PEPTIDES PRESENTED BY HLA CLASS II MOLECULES
; FILE REFERENCE: L0461/7052
; CURRENT APPLICATION NUMBER: US/09/166,448
; CURRENT FILING DATE: 1998-10-05
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 79
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-166-448-79

Query Match 100.0%; Score 61; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLMMWITOCFL 11
DB 1 SLMMWITOCFL 11

RESULT 6
US-09-705-160-42
; Sequence 42, Application US/09705160
; Patent No. 6387630
; GENERAL INFORMATION:
; APPLICANT: Van Baren, Nicolas
; APPLICANT: Brasseur, Francis
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: METHOD FOR DIAGNOSING LEUKEMIA BY DETERMINING
; FILE REFERENCE: LUD 5527.3-JEL/VAS
; CURRENT APPLICATION NUMBER: US/09/705,160
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: US 09/183,931
; PRIOR FILING DATE: 1998 - 10 - 30
; NUMBER OF SEQ ID NOS: 44
; SEQ ID NO 42
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION:
US-09-705-160-42

Query Match 100.0%; Score 61; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLMMWITOCFL 11
DB 1 SLMMWITOCFL 11

DB 1 SLMMWITOCFL 11

RESULT 7
US-09-440-621-1
; Sequence 1, Application US/09440621
; Patent No. 6417165
; GENERAL INFORMATION:
; APPLICANT: Valmori, Daniela
; APPLICANT: Cerottini, Jean-Charles
; APPLICANT: Romero, Pedro
; TITLE OF INVENTION: Derivatives of NY-ESO-1 Peptides and Uses Thereof
; FILE REFERENCE: LUD 5629
; CURRENT APPLICATION NUMBER: US/09/440,621
; CURRENT FILING DATE: 1999-11-15
; EARLIER APPLICATION NUMBER: US 09/165,546
; EARLIER FILING DATE: 1998-10-02
; EARLIER APPLICATION NUMBER: US 09/062,422
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: US 08/725,162
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 1
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-440-621-1

Query Match 100.0%; Score 61; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLMMWITOCFL 11
DB 1 SLMMWITOCFL 11

RESULT 8
US-09-697-884-79
; Sequence 79, Application US/09697884
; Patent No. 6426217
; GENERAL INFORMATION:
; APPLICANT: Chaux, Pascal
; APPLICANT: Vantomme, Valérie
; APPLICANT: Stroobant, Vincent
; APPLICANT: Boon-Falleur, Thierry
; APPLICANT: Van der Bruggen, Pierre
; APPLICANT: Thielemans, Kris
; APPLICANT: Cortbals, Jurgen
; TITLE OF INVENTION: MAGE-3 PEPTIDES PRESENTED BY HLA CLASS II MOLECULES
; FILE REFERENCE: L0461/7052
; CURRENT APPLICATION NUMBER: US/09/697,884
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: 09/166,448
; PRIOR FILING DATE: 1998-10-05
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 79
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-697-884-79

Query Match 100.0%; Score 61; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLMMWITOCFL 11
DB 1 SLMMWITOCFL 11

ZIP: 10103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/359,503
FILING DATE: July 23, 1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/165,546
FILING DATE: October 2, 1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/062,422
FILING DATE: April 17, 1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/937,263
FILING DATE: September 15, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/752,182
FILING DATE: 03-October-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6251603man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5466.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3000
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-359-503-4

Query Match 100.0%; Score 61; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLIMWITQCF 11
DB 1 SLIMWITQCF 11

RESULT 3
US-09-062-422-4
Sequence 4, Application US/09062422
Patent No. 6252052
GENERAL INFORMATION:
APPLICANT: Stockert, Elisabeth; Jager, Elke;
APPLICANT: Chen, Yao-Tseng; Scanlan, Matthew; Knuth, Alexander;
APPLICANT: Old, Lloyd J.
TITLE OF INVENTION: ANTIBODIES WHICH BIND TO NY-ESO-1
TITLE OF INVENTION: CANCER ASSOCIATED PROTEINS, USBS THEREOF, TRUNCATED
TITLE OF INVENTION: FORS OF NY-ESO-1, AND HLA BINDING PEPTIDES DERIVED
TITLE OF INVENTION: THEREFROM
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/062,422
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/937,263
FILING DATE: September 15, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/752,182
FILING DATE: 03-October-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6252052man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5466.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-062-422-4

Query Match 100.0%; Score 61; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLIMWITQCF 11
DB 1 SLIMWITQCF 11

RESULT 4
US-08-937-263B-4
Sequence 4, Application US/08937263B
Patent No. 6274145
GENERAL INFORMATION:
APPLICANT: Chen, Yao-Tseng; Scanlan, Matthew;
APPLICANT: Gure, Ali; Old, Lloyd J.; Jager, Elke;
APPLICANT: Alexander, Knuth; Drifflout, Jan W.
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE
TITLE OF INVENTION: ENCODING CANCER ASSOCIATED ANTIGEN, THE ANTIGEN
TITLE OF INVENTION: ITSELF, AND USBS THEREOF
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski, L.L.P.
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/937,263B
FILING DATE: September 15, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/725,182
FILING DATE: October 3, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Sinn, Eric; Patent Agent
REGISTRATION NUMBER: 40,177
REFERENCE/DOCKET NUMBER: LUD 5466.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3000
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-937-263B-4

Fri Feb 13 14:37:37 2004

us-10-023-182-4.ra1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model 1

Run on: February 10, 2004, 17:14:45 ; Search time 14.4138 Seconds
(without alignments)
32.290 Million cell updates/sec

Title: US-10-023-182-4
Perfect score: 61
Sequence: 1 SLIMWITOCFL 11

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/6C.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/6D.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	100.0	11	3	US-09-183-931-42
2	61	100.0	11	3	US-09-359-503-4
3	61	100.0	11	3	US-09-062-422-4
4	61	100.0	11	3	US-08-937-263B-4
5	61	100.0	11	3	US-09-166-448-79
6	61	100.0	11	4	US-09-705-160-42
7	61	100.0	11	4	US-09-440-621-1
8	61	100.0	11	4	US-09-697-884-79
9	61	100.0	11	4	US-09-751-798-4
10	61	100.0	11	4	US-09-344-040C-130
11	61	100.0	11	4	US-09-574-749B-33
12	61	100.0	180	2	US-08-791-495-7
13	61	100.0	180	2	US-08-791-495-9
14	61	100.0	180	2	US-08-937-263B-8
15	61	100.0	180	2	US-08-937-263B-8
16	57	96.9	10	3	US-09-359-503-7
17	53	93.4	9	4	US-09-344-040C-123
18	51	83.6	9	3	US-09-183-931-43
19	51	83.6	9	3	US-09-359-503-5
20	51	83.6	9	3	US-09-062-422-5
21	51	83.6	9	3	US-08-937-263B-5
22	51	83.6	9	3	US-09-166-448-80
23	51	83.6	9	4	US-09-705-160-43
24	51	83.6	9	4	US-09-440-621-2
25	51	83.6	9	4	US-09-697-884-80
26	51	83.6	9	4	US-09-697-884-80
27	51	83.6	9	4	US-09-697-884-80

28	51	83.6	9	4	US-09-344-040C-131
29	51	83.6	9	4	US-09-574-749B-34
30	42	68.9	9	4	US-09-440-621-6
31	42	68.9	9	4	US-09-440-621-7
32	42	68.9	9	4	US-09-440-621-8
33	42	68.9	9	4	US-09-440-621-10
34	37	60.7	9	3	US-09-183-931-44
35	37	60.7	9	3	US-09-359-503-6
36	37	60.7	9	3	US-09-062-422-6
37	37	60.7	9	3	US-08-937-263B-6
38	37	60.7	9	3	US-09-166-448-81
39	37	60.7	9	4	US-09-705-160-44
40	37	60.7	9	4	US-09-440-621-3
41	37	60.7	9	4	US-09-697-884-81
42	37	60.7	9	4	US-09-751-798-6
43	37	60.7	9	4	US-09-344-040C-122
44	37	60.7	9	4	US-09-344-040C-132
45	37	60.7	9	4	US-09-574-749B-35

ALIGNMENTS

RESULT 1
US-09-183-931-42
; Sequence 42, Application US/09183931C
; Patent No. 6210886
; GENERAL INFORMATION:
; APPLICANT: Van Baren, Nicolas
; APPLICANT: Brasseur, Francis
; APPLICANT: Boon-Falheur, Thierry
; TITLE OF INVENTION: METHOD FOR DIAGNOSING LEUKEMIA BY DETERMINING
; TITLE OF INVENTION: TUMOR REJECTION ANTIGEN PRECURSORS
; FILE REFERENCE: LUD 5527.1-JEL/ES
; CURRENT APPLICATION NUMBER: US/09/183,931C
; EARLIER FILING DATE: 2000-02-28
; EARLIER APPLICATION NUMBER: US 09/018,422
; NUMBER OF SEQ ID NOS: 44
; SEQ ID NO 42
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-09-183-931-42

Query Match 100.0%; Score 61; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 SLIMWITOCFL 11
DB 1 SLIMWITOCFL 11

RESULT 2
US-09-359-503-4
; Sequence 4, Application US/09359503
; Patent No. 6251603
; GENERAL INFORMATION:
; APPLICANT: Jager, Elke; Stockert, Elisabeth; Old, Lloyd J;
; APPLICANT: Knuth, Alexander
; TITLE OF INVENTION: Method for Determining Status of A
; TITLE OF INVENTION: Cancerous Condition By Determining Antibodies
; TITLE OF INVENTION: a Patient Sample
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski, L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA

Fri Feb 13 14:37:41 2004

us-10-023-182-6.ral

Page 6

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 144 Kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/937,263B
FILING DATE: September 15, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/725,182
FILING DATE: October 3, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Sinn, Eric, Patent Agent
REGISTRATION NUMBER: 40,177
REFERENCE/DOCKET NUMBER: LUD 5466.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3000
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 180
TYPE: amino acid
TOPOLOGY: linear
US-08-937-263B-8

Query Match

100.0%; Score 46; DB 3; Length 180;

Best Local Similarity 100.0%; Pred. No. 0.66; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLSLMMIT 9
Db 155 QLSLMMIT 163

Search completed: February 10, 2004, 17:20:32
Job time: 12.7931 secs

PRIOR APPLICATION NUMBER: PCT/US99/26795
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: US 09/524,749
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 58
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 35
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Homo Sapiens source
US-09-574-749B-35

Query Match 100.0%; Score 46; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QLSLWMT 9
155 QLSLWMT 163

RESULT 13
US-08-791-495-7
Sequence 7, Application US/08791495
Patent No. 5811519
GENERAL INFORMATION:
APPLICANT: Leth, Bernard
APPLICANT: Lucas, Sophie
APPLICANT: De Smet, Charles
APPLICANT: Godelaine, Daniele
APPLICANT: Boon-Falleur, Thierry
TITLE OF INVENTION: LL-1 TUMOR SPECIFIC GENES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/791,495
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Van Amsterdam, John R.
REGISTRATION NUMBER: 40,212
REFERENCE/DOCKET NUMBER: L0461/7005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 180 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-791-495-7

Query Match 100.0%; Score 46; DB 2; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QLSLWMT 9
155 QLSLWMT 163

RESULT 14
US-08-791-495-9
Sequence 9, Application US/08791495
Patent No. 5811519
GENERAL INFORMATION:
APPLICANT: Leth, Bernard
APPLICANT: Lucas, Sophie
APPLICANT: De Smet, Charles
APPLICANT: Godelaine, Daniele
APPLICANT: Boon-Falleur, Thierry
TITLE OF INVENTION: LL-1 TUMOR SPECIFIC GENES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/791,495
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Van Amsterdam, John R.
REGISTRATION NUMBER: 40,212
REFERENCE/DOCKET NUMBER: L0461/7005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 180 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-791-495-9

Query Match 100.0%; Score 46; DB 2; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QLSLWMT 9
155 QLSLWMT 163

RESULT 15
US-08-937-263B-8
Sequence 8, Application US/08937263B
Patent No. 6274145
GENERAL INFORMATION:
APPLICANT: Chen, Yao-Tseng; Scanlan, Matthew;
APPLICANT: Gure, Ali; Old, Lloyd J.; Jager, Elke;
APPLICANT: Alexander, Knuth; Dijkhout, Jan W.
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE
TITLE OF INVENTION: ENCODING CANCER ASSOCIATED ANTIGEN, THE ANTIGEN
TITLE OF INVENTION: ITSELF, AND USES THEREOF
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski, L.L.P.
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10103

```
RESULT 9
US-09-751-798-6
; Sequence 6, Application US/09751798
; Patent No. 6525177
; GENERAL INFORMATION:
; APPLICANT: Stockert, Elisabeth; Jager, Elke;
; APPLICANT: Chen, Yao-tseng; Scanlan, Matthew;
; APPLICANT: Knuth, Alexander; Old, Lloyd J.
; TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer
; TITLE OF INVENTION: Associated Proteins, Uses Thereof,
; TITLE OF INVENTION: Truncated Forms of NY-ESO-1, and HLA
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Fulbright & Jaworski, L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/751,798
; FILING DATE:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 09/062,422
; FILING DATE:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/725,182
; FILING DATE: October 3, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6525177man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5466.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3168
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-751-798-6
Query Match 100.0%; Score 46; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QLSLWMIT 9
Db 1 QLSLWMIT 9
RESULT 10
US-09-344-040C-122
; Sequence 122, Application US/09344040C
; Patent No. 6548064
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfrendschuh, Michael
; APPLICANT: Ramensee, Hans Georg
; APPLICANT: Stevanovic, Stefan
; TITLE OF INVENTION: Method for Determining Presence of Cancer in a Sample By Determin
; TITLE OF INVENTION: Expression of an SSX Gene, Peptides Derived From Said SSX Gene a
; FILE REFERENCE: LUD 5556.1
; CURRENT APPLICATION NUMBER: US/09/344,040C
; CURRENT FILING DATE: 1999-06-25
```

```
;; PRIOR APPLICATION NUMBER: US 09/105,839
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: US 08/851,130
;; PRIOR FILING DATE: 1997-05-05
;; NUMBER OF SEQ ID NOS: 132
;; SEQ ID NO 122
;; LENGTH: 9
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-344-040C-122
Query Match 100.0%; Score 46; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QLSLWMIT 9
Db 1 QLSLWMIT 9
RESULT 11
US-09-344-040C-132
; Sequence 132, Application US/09344040C
; Patent No. 6548064
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfrendschuh, Michael
; APPLICANT: Ramensee, Hans Georg
; APPLICANT: Stevanovic, Stefan
; TITLE OF INVENTION: Method for Determining Presence of Cancer in a Sample By Determin
; TITLE OF INVENTION: Expression of an SSX Gene, Peptides Derived From Said SSX Gene
; FILE REFERENCE: LUD 5556.1
; CURRENT APPLICATION NUMBER: US/09/344,040C
; CURRENT FILING DATE: 1999-06-25
; PRIORITY APPLICATION NUMBER: US 09/105,839
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: US 08/851,130
; PRIOR FILING DATE: 1997-05-05
; NUMBER OF SEQ ID NOS: 132
; SEQ ID NO 132
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-344-040C-132
Query Match 100.0%; Score 46; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QLSLWMIT 9
Db 1 QLSLWMIT 9
RESULT 12
US-09-574-749B-35
; Sequence 35, Application US/09574749B
; Patent No. 6548299
; GENERAL INFORMATION:
; APPLICANT: ROSENZWEIG, Michael
; APPLICANT: PYKEET, Mark J.
; APPLICANT: SCADEN, David T.
; APPLICANT: POZNANSKY, Mark C.
; TITLE OF INVENTION: LYMPHOID TISSUE-SPECIFIC CELL PRODUCTION
; TITLE OF INVENTION: DEVICES
; FILE REFERENCE: C1005/7012/Ka/ERG
; CURRENT APPLICATION NUMBER: US/09/574,749B
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US 60/107,972
; PRIOR FILING DATE: 1998-11-12
```

Query Match 100.0%; Score 46; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLSLMMIT 9
Db 1 QLSLMMIT 9

RESULT 5
US-09-166-448-81
; Sequence 81, Application US/09166448
; Patent No. 6291430
; GENERAL INFORMATION:
; APPLICANT: Chauv, Pascal
; APPLICANT: Vancomme, Valrie
; APPLICANT: Stroobant, Vincent
; APPLICANT: Boon-Falleur, Thierry
; APPLICANT: Van der Bruggen, Pierre
; APPLICANT: Thielemans, Kris
; APPLICANT: Corbals, Jurgen
; TITLE OF INVENTION: MAGE-3 PEPTIDES PRESENTED BY HLA CLASS II MOLECULES
; FILE REFERENCE: L0461/7052
; CURRENT APPLICATION NUMBER: US/09/166,448
; CURRENT FILING DATE: 1998-10-05
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 81
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-166-448-81

Query Match 100.0%; Score 46; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLSLMMIT 9
Db 1 QLSLMMIT 9

RESULT 6
US-09-705-160-44
; Sequence 44, Application US/09705160
; Patent No. 6387630
; GENERAL INFORMATION:
; APPLICANT: Van Baren, Nicolas
; APPLICANT: Brasseur, Francis
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: METHOD FOR DIAGNOSING LEUKEMIA BY DETERMINING
; FILE REFERENCE: LUD 5527.3-JEL/MAS
; CURRENT APPLICATION NUMBER: US/09/705,160
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: US 09/183,931
; PRIOR FILING DATE: 1998-10-30
; NUMBER OF SEQ ID NOS: 44
; SEQ ID NO 44
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION:
US-09-705-160-44

Query Match 100.0%; Score 46; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLSLMMIT 9

Db 1 QLSLMMIT 9

RESULT 7
US-09-440-621-3
; Sequence 3, Application US/09440621
; Patent No. 6417165
; GENERAL INFORMATION:
; APPLICANT: Valmori, Daniela
; APPLICANT: Cerottini, Jean-Charles
; APPLICANT: Romero, Pedro
; TITLE OF INVENTION: Derivatives of NY-ESO-1 Peptides and Uses Thereof
; FILE REFERENCE: LUD 5629
; CURRENT APPLICATION NUMBER: US/09/440,621
; CURRENT FILING DATE: 1999-11-15
; EARLIER APPLICATION NUMBER: US 09/165,546
; EARLIER FILING DATE: 1998-10-02
; EARLIER APPLICATION NUMBER: US 09/062,422
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: US 08/725,162
; EARLIER FILING DATE: 1996-10-03
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 3
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
US-09-440-621-3

Query Match 100.0%; Score 46; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLSLMMIT 9
Db 1 QLSLMMIT 9

RESULT 8
US-09-697-884-81
; Sequence 81, Application US/09697884
; Patent No. 6426217
; GENERAL INFORMATION:
; APPLICANT: Chauv, Pascal
; APPLICANT: Vancomme, Valrie
; APPLICANT: Stroobant, Vincent
; APPLICANT: Boon-Falleur, Thierry
; APPLICANT: Van der Bruggen, Pierre
; APPLICANT: Thielemans, Kris
; APPLICANT: Corbals, Jurgen
; TITLE OF INVENTION: MAGE-3 PEPTIDES PRESENTED BY HLA CLASS II MOLECULES
; FILE REFERENCE: L0461/7052
; CURRENT APPLICATION NUMBER: US/09/697,884
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: 09/166,448
; PRIOR FILING DATE: 1998-10-05
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 81
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-697-884-81

Query Match 100.0%; Score 46; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLSLMMIT 9
Db 1 QLSLMMIT 9

ZIP: 10103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/359,503
FILING DATE: July 23, 1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/165,546
FILING DATE: October 2, 1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/062,422
FILING DATE: April 17, 1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/937,263
FILING DATE: September 15, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/752,182
FILING DATE: 03-October-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6251603man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5466.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3000
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-359-503-6
Query Match 100.0%; Score 46; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QLSLMMIT 9
Db 1 QLSLMMIT 9
RESULT 3
US-09-062-422-6
Sequence 6, Application US/09062422
Patent No. 6252052
GENERAL INFORMATION:
APPLICANT: Stockert, Elisabeth; Jager, Elke;
APPLICANT: Chen, Yao-Tseng; Scanlan, Matthew; Knuth, Alexander;
APPLICANT: Old, Lloyd J.
TITLE OF INVENTION: ANTIBODIES WHICH BIND TO NY-ESO-1
TITLE OF INVENTION: CANCER ASSOCIATED PROTEINS, USES THEREOF, TRUNCATED
TITLE OF INVENTION: FORMS OF NY-ESO-1, AND HLA BINDING PEPTIDES DERIVED
TITLE OF INVENTION: THEREFROM
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felte & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/062,422
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/937,263
FILING DATE: September 15, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/752,182
FILING DATE: 03-October-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6252052man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5466.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-062-422-6
Query Match 100.0%; Score 46; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QLSLMMIT 9
Db 1 QLSLMMIT 9
RESULT 4
US-08-937-263B-6
Sequence 6, Application US/08937263B
Patent No. 6274145
GENERAL INFORMATION:
APPLICANT: Chen, Yao-Tseng; Scanlan, Matthew;
APPLICANT: Gure, Ali; Old, Lloyd J.; Jager, Elke;
APPLICANT: Alexander, Knuth; Drifhout, Jan W.
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE
TITLE OF INVENTION: ENCODING CANCER ASSOCIATED ANTIGEN, THE ANTIGEN
TITLE OF INVENTION: ITSELF, AND USES THEREOF
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski, L.L.P.
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/937,263B
FILING DATE: September 15, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/725,182
FILING DATE: October 3, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Sinn, Eric, Patent Agent
REGISTRATION NUMBER: 40,177
REFERENCE/DOCKET NUMBER: LUD 5466.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3000
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-937-263B-6

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 17:14:45 ; Search time 11.7931 Seconds
(without alignments)
32.290 Million cell updates/sec

Title: US-10-023-182-6
Perfect score: 46
Sequence: 1 QLSLWMIT 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	9	US-09-183-931-44	Sequence 44, Appl
2	46	100.0	9	US-09-359-503-6	Sequence 6, Appl
3	46	100.0	9	US-09-063-422-6	Sequence 6, Appl
4	46	100.0	9	US-08-937-263B-6	Sequence 6, Appl
5	46	100.0	9	US-09-166-448-81	Sequence 81, Appl
6	46	100.0	9	US-09-705-160-44	Sequence 44, Appl
7	46	100.0	9	US-09-440-621-3	Sequence 3, Appl
8	46	100.0	9	US-09-697-884-81	Sequence 8, Appl
9	46	100.0	9	US-09-751-798-6	Sequence 6, Appl
10	46	100.0	9	US-09-344-040C-122	Sequence 122, App
11	46	100.0	9	US-09-344-040C-132	Sequence 132, App
12	46	100.0	9	US-09-574-749B-35	Sequence 35, Appl
13	46	100.0	180	US-08-791-495-7	Sequence 7, Appl
14	46	100.0	180	US-08-791-495-9	Sequence 9, Appl
15	46	100.0	180	US-08-937-263B-8	Sequence 8, Appl
16	46	100.0	180	US-09-751-798-8	Sequence 8, Appl
17	37	80.4	9	US-09-183-931-43	Sequence 43, Appl
18	37	80.4	9	US-09-359-503-5	Sequence 5, Appl
19	37	80.4	9	US-09-063-422-5	Sequence 5, Appl
20	37	80.4	9	US-08-937-263B-5	Sequence 5, Appl
21	37	80.4	9	US-09-166-448-80	Sequence 80, Appl
22	37	80.4	9	US-09-705-160-43	Sequence 43, Appl
23	37	80.4	9	US-09-440-621-2	Sequence 2, Appl
24	37	80.4	9	US-09-440-621-6	Sequence 6, Appl
25	37	80.4	9	US-09-440-621-7	Sequence 7, Appl
26	37	80.4	9	US-09-440-621-8	Sequence 8, Appl
27	37	80.4	9	US-09-440-621-10	Sequence 10, Appl

28	37	80.4	9	US-09-697-884-80	Sequence 80, Appl
29	37	80.4	9	US-09-670-456A-1	Sequence 1, Appl
30	37	80.4	9	US-09-751-798-5	Sequence 5, Appl
31	37	80.4	9	US-09-344-040C-131	Sequence 131, Appl
32	37	80.4	9	US-09-574-749B-34	Sequence 34, Appl
33	37	80.4	11	US-09-183-931-42	Sequence 42, Appl
34	37	80.4	11	US-09-359-503-4	Sequence 4, Appl
35	37	80.4	11	US-09-062-422-4	Sequence 4, Appl
36	37	80.4	11	US-08-937-263B-4	Sequence 4, Appl
37	37	80.4	11	US-09-166-448-79	Sequence 79, Appl
38	37	80.4	11	US-09-705-160-42	Sequence 42, Appl
39	37	80.4	11	US-09-440-621-1	Sequence 1, Appl
40	37	80.4	11	US-09-697-884-79	Sequence 79, Appl
41	37	80.4	11	US-09-751-798-4	Sequence 4, Appl
42	37	80.4	11	US-09-344-040C-130	Sequence 130, Appl
43	37	80.4	11	US-09-574-749B-33	Sequence 33, Appl
44	35	76.1	4	US-09-252-991A-30324	Sequence 30324, A
45	33	71.7	6	US-09-359-503-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-183-931-44
Sequence 44, Application US/09183931C
Patent No. 6210866
GENERAL INFORMATION:
APPLICANT: Van Baren, Nicolas
APPLICANT: Brasseur, Francis
APPLICANT: Boon-Falheur, Thierry
TITLE OF INVENTION: METHOD FOR DIAGNOSING LEUKEMIA BY DETERMINING
TITLE OF INVENTION: TUMOR REJECTION ANTIGEN PRECURSORS
FILE REFERENCE: IUD 5527.1-JEU/ES
CURRENT APPLICATION NUMBER: US/09/183,931C
CURRENT FILING DATE: 2000-02-28
EARLIER APPLICATION NUMBER: US 09/018,422
EARLIER FILING DATE: 1998 - 02 - 04
NUMBER OF SEQ ID NOS: 44
SEQ ID NO 44
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION:
US-09-183-931-44

Query Match 100.0%; Score 46; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLSLWMIT 9
Db 1 QLSLWMIT 9

RESULT 2
US-09-359-503-6
Sequence 6, Application US/09359503
Patent No. 6251603
GENERAL INFORMATION:
APPLICANT: Jager, Elke; Stockert, Elisabeth; Old, Lloyd J;
TITLE OF INVENTION: Method for Determining Status of A
TITLE OF INVENTION: Cancerous Condition By Determining Antibodies
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: Fulbright & Jaworski, L.L.P.
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA

```

XX
DT 20-APR-2001 (first entry)
XX
DE Exemplary antigen characteristic of tumours and derived from NY-ESO-1.
XX
KW MAGE-A1; HLA; human leukocyte antigen; CD4+ T lymphocyte; cancer;
XX MAGE-A1 HLA class II-binding protein; vaccine.
XX
OS Homo sapiens.
XX
PN WO200078806-A1.
XX
PD 28-DEC-2000.
XX
PF 14-JUN-2000; 2000WO-DS16287.
XX
PR 18-JUN-1999; 99US-0336091.
XX
PA (LUDM-) LUDMIG INST CANCER RES.
XX
PI Van Snick J, Lethe B, Chaux P, Boon-Falleur T, Van Der Bruggen P;
XX WPI; 2001-102698/11.
XX
PT Novel MAGE-A1 human leukocyte antigen class II peptides which bind to
PT and are presented to the class II molecules, useful for inducing immune
PT response and treating cancers characterized by expression of MAGE-A1 -
XX
PS Disclosure; Page 32; 78pp; English.
XX
CC AAB31302-59 represent exemplary antigens which are characteristic
CC of tumours. They can be used to enhance the immune response of vaccines
CC comprising peptides derived from human MAGE-A1 HLA (human leukocyte
CC antigen) class II-binding protein. Peptides derived from the MAGE-A1 HLA
CC binding protein stimulate the activity and proliferation of CD4+ T
CC lymphocytes. The MAGE-A1 HLA binding protein is useful as a diagnostic
CC agent for diagnosing a disorder characterized by expression of MAGE-A1.
CC The protein is used for treating a disorder characterized by expression
CC of MAGE-A1 such as cancers e.g. melanoma, squamous cell carcinomas,
CC colorectal carcinomas, osteosarcomas, and lymphocytic leukemias.
CC Peptides derived from the MAGE-A1 HLA binding protein are useful in the
CC production of anti-tumour vaccines.
XX
SQ Sequence 9 AA;

```

Query Match 100.0%; Score 51; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 SLIMWITQC 9
   |||||
   |||||
   |||||
   |||||
   |||||
Db 1 SLIMWITQC 9

```

Search completed: February 10, 2004, 17:17:29
 Job time : 32.5862 secs

CC is useful for screening for a cancerous condition, which involves
 CC contacting a subject sample to a cell line transacted with the
 CC immunoreactive cell (helper T cell), where interaction is indicative of
 CC cancer. In addition, a sample from a patient (for example, a body fluid
 CC or tissue) can be monitored for the amount of the complex present in the
 CC bloodstream. This is useful for determining regression, progression or
 CC onset of a cancerous condition. The method involves contacting the sample
 CC with a radioactive labeled or enzyme labelled monoclonal antibody which
 CC specifically binds with the complex.

XX Sequence 9 AA;

Query Match 100.0%; Score 51; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLMMWITQC 9
 DB 1 SLMMWITQC 9

RESULT 13

AAB82017
 ID AAB82017 standard; Peptide; 9 AA.

XX AAB82017;

DT 12-JUN-2001 (first entry)

DE HLA- binding peptide derived from NY-ESO-1.

XX Multiple myeloma; tumour rejection antigen precursor; MAGE; BAGE; GAGE;
 KW LAGE; NY-ESO-1; PRAME; DAGG; human; HLA.

XX Homo sapiens.

XX US6210886-B1.

XX 03-APR-2001.

XX 30-OCT-1998; 98US-0183931.

XX 04-FEB-1998; 98US-0018422.

XX (LUDM-) LUDWIG INST CANCER RES.

XX Van Baren N, Brasseur F, Boon-Falleur T;

XX WPI; 2001-289628/30.

XX Detecting multiple myeloma in a patient, comprises contacting a nucleic
 PT acid containing sample taken from bone marrow or blood with a
 PT hybridization probe specific for a tumor rejection antigen precursor -

XX Example 3; Column 11; 16pp; English.

XX The present invention relates to a method for detecting multiple myeloma.
 CC The method comprises contacting a nucleic acid containing a sample taken
 CC from a bone marrow or blood of a patient, with a hybridisation probe
 CC specific for a tumour rejection antigen precursor. Tumour rejection
 CC antigen precursors used in the present invention are the MAGE family,
 CC BAGE, GAGE, LAGE, NY-ESO-1 and PRAME (previously referred to as DAGG).
 CC Expression of the tumour rejection antigen precursor indicates possible
 CC multiple myeloma in the patient. The method can also be used for
 CC monitoring the disease progress and course of therapeutic regime. The
 CC present sequence is a peptide derived from a tumour rejection antigen
 CC precursor, which was used in the method of the present invention.

XX Sequence 9 AA;

Query Match 100.0%; Score 51; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLMMWITQC 9
 DB 1 SLMMWITQC 9

RESULT 14

AAB69948
 ID AAB69948 standard; Peptide; 9 AA.

XX AAB69948;

DT 27-APR-2001 (first entry)

DE Human NY-ESO-1 CTL stimulating peptide #2.

XX Human; NY-ESO-1; HLA; human leukocyte antigen; CTL; cytotoxic T cell;
 KW HLA-A2; HLA-DR53; melanoma; adenocarcinoma; bladder carcinoma;
 KW non-small cell lung carcinoma; tumour status determination.

XX Homo sapiens.

XX WO200107917-A1.

XX 01-FEB-2001.

XX 14-JUL-2000; 2000WO-US19220.

XX 23-JUL-1999; 99US-0359503.

XX (LUDM-) LUDWIG INST CANCER RES.

XX (SLOK) SLOAN KETTERING INST CANCER RES.

XX (CORR) CORNELL RES FOUND INC.

XX Jager E, Stockert E, Old LJ, Knuth A, Chen Y, Scanlan M;

XX WPI; 2001-182822/18.

XX Method useful for determining the status (e.g. progression, regression
 PT or stability of the disease) of a cancerous condition, involves
 PT determining the levels of NY-ESO-1 specific antibodies in a sample
 PT taken from a patient -

XX Example 13; Page 24; 50pp; English.

XX The present sequence is given in a specification relating to a method
 CC for determining the status of a cancerous condition in a patient
 CC with a tumour that expresses NY-ESO-1. The method comprises assaying a
 CC sample taken from the patient for antibodies that specifically bind to
 CC the NY-ESO-1 and comparing the value obtained to a prior value obtained
 CC from assay of a prior sample taken from the patient. Any difference
 CC between the values is indicative of a change in status of the cancerous
 CC condition. The method is useful for determining whether a cancerous
 CC condition is progressing, regressing or remaining stable, in particular
 CC in patients receiving treatment for a melanoma, adenocarcinoma,
 CC non-small cell lung carcinoma or bladder carcinoma.

XX Sequence 9 AA;

Query Match 100.0%; Score 51; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLMMWITQC 9
 DB 1 SLMMWITQC 9

RESULT 15

AAB31329
 ID AAB31329 standard; peptide; 9 AA.

XX AAB31329;

XX The invention provides NY-ESO-1 peptide derivatives which bind to human
 CC leukocyte antigen (HLA) molecules and provokes lysis by cytolytic T cells
 CC (CTLs). The NY-ESO-1 nonapeptide is of formula SLIWMITQC, where X is an
 CC amino acid having an uncharged polar side chain. The NY-ESO-1 peptide
 CC derivatives are useful for determining if a cell contains an HLA-A2
 CC molecule on its surface, by contacting a sample containing the cell with
 CC the peptide or its derivative, and determining binding between them,
 CC where the binding is indicative of HLA-A2 on the surface of the cell.
 CC The NY-ESO-1 peptides and analogues are useful therapeutically, for
 CC administration to a patient who is HLA-A2 positive and expresses NY-ESO-1
 CC in connection with the pathology, as well as diagnostically, i.e. to
 CC determine if HLA-A2 positive cells are present, or if relevant CTLs are
 CC present. They are also useful for determining the presence of CTLs in a
 CC sample. The peptides are useful as T-cell sorters, when incorporated into
 CC immune tetramers. The present sequence represents a NY-ESO-1 peptide that
 CC can bind to HLA-A2 molecule.

Sequence 9 AA;

Query Match 100.0%; Score 51; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLIWMITQC 9
 DB 1 SLIWMITQC 9

RESULT 11
 ID AAE02120 standard; peptide; 9 AA.

AAE02120;

31-JUL-2001 (first entry)

NY-ESO-1 human leukocyte antigen-A2-binding peptide #2.

Human; cytotoxic; immunogen; NY-ESO-1; human leukocyte antigen;
 HLA; C8; cytotoxic T lymphocyte; cancer; carcinoma; melanoma; myeloma;
 Brain tumour; sarcoma; vaccine; gene therapy.

Homo sapiens.

MO200129220-A2.

26-APR-2001.

19-OCT-2000; 2000MO-US28852.

19-OCT-1999; 99US-0160374.

01-FEB-2000; 2000US-0179570.

(LUDW-) LUDWIG INST CANCER RES.

Heidecker L, Van Den Eynde B, Boon-Falleur T, Brasseur F;

WPI; 2001-328498/34.

New antigenic peptides derived from MAGE-A12 polypeptides, useful for
 diagnosis and treatment of cancer, such as bladder, lung, breast,
 brain, prostate and renal carcinomas -

Disclosure; Page 21; 69pp; English.

The patent discloses antigenic peptides derived from MAGE-A12
 protein and presented by human leukocyte antigens (HLAs). These
 antigenic peptides when presented by an antigen presenting cell
 having a HLA class I molecule, effectively induce the activation
 and proliferation of CD8+ cytotoxic T lymphocytes (CTLs). MAGE-A12
 is useful for treating a subject having a disorder characterised by
 expression of MAGE-A12. The protein microarray comprising MAGE-A12

CC is useful for diagnosing a disorder, especially cancer, by determining
 CC the binding of an antibody, T lymphocytes or a HLA molecule isolated
 CC from the subject suspected of having the disorder characterised by the
 CC expression of MAGE-A12. MAGE-A12 is useful for treating cancers,
 CC including bladder carcinomas, melanomas, oesophageal, lung, head and
 CC neck, breast, colorectal carcinomas, myelomas, brain tumours, sarcomas,
 CC prostate and renal carcinomas and to produce antibodies. MAGE-A12
 CC antibodies are useful for diagnosing disorders characterised by
 CC expression of MAGE-A12 immunogenic polypeptide. These MAGE-A12 peptides
 CC are used as vaccines. They are also used in gene therapy.
 CC The present sequence is an antigenic peptide derived from NY-ESO-1. This
 CC peptide which is characteristic of tumours is presented by HLA-A2 MHC
 CC (major histocompatibility complex) and is recognised by CTLs.

Sequence 9 AA;

Query Match 100.0%; Score 51; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLIWMITQC 9
 DB 1 SLIWMITQC 9

RESULT 12
 ID AAU01537 standard; Peptide; 9 AA.

AAU01537;

18-JUL-2001 (first entry)

Cytolytic T cell line stimulator peptide #2.

NY-ESO-1; human; tumour rejection antigen precursor; SSX-2; MHC Class II;
 major histocompatibility complex; helper T cell; HLA-DR; cancer;
 human leukocyte antigen-determining region; disease progression;
 disease regression; disease onset; body tissue; body fluid; enzyme label;
 radioactive label; monoclonal antibody; cytolytic T cell line.

Homo sapiens.

MO200123560-A2.

05-APR-2001.

26-SEP-2000; 2000MO-US26411.

29-SEP-1999; 99US-0408036.

(LUDW-) LUDWIG INST CANCER RES.

Tureci O, Sahin U, Pfreundschuh M;

WPI; 2001-266156/27.

Polypeptides binding to major histocompatibility complex class II human
 leukocyte antigen-determining region molecule having amino acid
 sequence found in tumour rejection antigen precursor used for
 stimulating proliferation of helper T cells -

Example 12; Page 17; 62pp; English.

The sequence represents a human NY-ESO-1 tumour rejection antigen
 precursor fragment which efficiently stimulates cytolytic T cell lines.
 CC NY-ESO-1 and SSX-2 polypeptides, or fragments of, bind to major
 CC histocompatibility complex (MHC) Class II molecules such as human
 CC leukocyte antigen-determining region (HLA-DR) molecules and stimulate
 CC proliferation of helper T cells. The peptides can be administered to an
 CC HLA-DR positive subject in order to stimulate the helper T cells. An MHC
 CC class II HLA-DR-NY-ESO-1/SSX-2 complex expressed on the surface of a cell
 CC or present in free form is useful for this stimulation. The nucleic acid


```
XX 22-FEB-2000; 2000US-0510635.
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Lethe B, Boon-Falleur T;
XX
XX WPI; 2001-550091/61.
XX
XX Genomic sequences of tumour associated antigen EY-ESO-1 (LAGE-2) useful
XX for diagnosing testicular tumours -
XX
XX Example 12; Page 24; 50pp; English.
XX
XX The present sequence represents a peptide which is derived from cancer
XX testis tumour antigen NY-ESO-1 (also called LAGE-2). The peptide that
XX stimulates cytolytic T cell lines (CTLs). NY-ESO-1 is a molecule that
XX is processed to at least one human leukocyte antigen (HLA) binding
XX peptide, which binds to Class I and Class II major histocompatibility
XX complex (MHC). NY-ESO-1 is expressed in tumour mRNA and in testis,
XX but not normal colon, kidney, liver or brain tissue. The presence of
XX level of expression of NY-ESO-1 may be assayed for the diagnosis of
XX cancer, especially testis tumours.
XX
XX
XX Sequence 9 AA;
SQ
Query Match 100.0%; Score 51; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLIMWITQC 9
DB 1 SLIMWITQC 9
RESULT 9
AAE06850
ID AAE06850 standard; peptide; 9 AA.
XX
XX AAE06850;
XX
XX 16-OCT-2001 (first entry)
XX
XX Human NY-ESO-1 antigenic peptide #2.
XX
XX MAGE antigenic peptide; Human leukocyte antigen; HLA-B35; HLA-B44;
XX tumour cell; immunostimulant; antigen presentation; cancer; melanoma;
XX CD8+ cytotoxic T lymphocyte; colorectal; prostate; gastric carcinoma;
XX myeloma; brain tumour; sarcoma; seminoma; ovarian tumour; cystostatic;
XX Gene therapy; tumour rejection antigen; TRA; human; NY-ESO-1;
XX MHC; major histocompatibility complex.
XX
XX Homo sapiens.
XX
XX WO200153833-A1.
XX
XX 26-JUL-2001.
XX
XX 19-JAN-2001; 2001MO-US02008.
XX
XX 20-JAN-2000; 2000US-0177242.
XX
XX 25-OCT-2000; 2000US-0243212.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Luiten R, Boon-Falleur T, Van Der Bruggen P, Stroobant V;
XX Demotte N, Schultz E;
XX
XX WPI; 2001-488724/53.
XX
XX Functional variants and isolated mimetics of a MAGE-A1 HLA-B35 or
XX HLA-B44 binding peptide, or of a MAGE-A3 HLA-B35 binding peptide, used
XX in diagnosis and treatment of a disorder characterized by expression of
```

```
PT MAGE-A1 or -A3 -
XX
XX Disclosure; Page 28; 103pp; English.
XX
XX The invention relates to functional variants and isolated mimetics of a
XX MAGE-A1 human leukocyte antigen (HLA)-B35 or HLA-B44 binding peptide,
XX or of a MAGE-A3 HLA-B35 binding peptide, identified by methods described
XX in the specification. MAGE genes encode tumour rejection antigens
XX (TRAs) presented to T lymphocytes by HLA-B35 and HLA-B44 molecules.
XX The MAGE antigenic peptide acts by binding to HLA molecules
XX on tumour cells and stimulating recognition of these cells and thus
XX signalling them to the immune system for destruction. The peptide when
XX presented by HLA molecule induces the activation and stimulation of
XX CD8+ cytotoxic T lymphocytes. The MAGE antigenic peptide is used to
XX treat and diagnose disorders characterised by expression of MAGE-A1
XX or -A3. Disorders include cancers e.g. melanomas, oesophageal, lung,
XX head and neck, breast, colorectal, prostate, renal, bladder,
XX hepatocellular, papillary thyroid and gastric carcinomas, myelomas,
XX brain tumours, sarcomas, seminomas, and ovarian tumours. The present
XX sequence is human NY-ESO-1 tumour associated antigenic peptide
XX presented by major histocompatibility complex (MHC) HLA-A2. The
XX antigenic peptide is used in combination with peptides of the
XX invention for inducing an immune response.
XX
XX
XX Sequence 9 AA;
SQ
Query Match 100.0%; Score 51; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLIMWITQC 9
DB 1 SLIMWITQC 9
RESULT 10
AAB85299
ID AAB85299 standard; peptide; 9 AA.
XX
XX AAB85299;
XX
XX 17-SEP-2001 (first entry)
XX
XX HLA-A2 binding NY-ESO-1 peptide #2.
XX
XX NY-ESO-1; human leukocyte antigen; HLA; lysis; cytolytic T cell; CTL;
XX HLA-A2; T-cell sorter; tumour; immune tetramer.
XX
XX Homo sapiens.
XX
XX WO200136453-A2.
XX
XX 25-MAY-2001.
XX
XX 08-NOV-2000; 2000WO-US42010.
XX
XX 15-NOV-1999; 99US-0440621.
XX
XX 25-FEB-2000; 2000US-0514036.
XX
XX 29-SEP-2000; 2000US-0676005.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX (UYOX-) UNIV OXFORD.
XX
XX Valmorì D, Cerottini J, Romero P, Cerundolo V;
XX
XX WPI; 2001-451454/48.
XX
XX Novel isolated NY-ESO-1 nonapeptide useful for determining if a cell
XX presents human leukocyte antigen-A2 molecule on its surface, binds to
XX human leukocyte antigen molecules and provokes lysis by cytolytic T
XX cells -
XX
XX Example 1; Page 4; 38pp; English.
XX
```

```

XX Homo sapiens.
OS
XX MO200000824-A1.
XX
XX 06-JAN-2000.
XX
XX 25-JUN-1999; 99WO-US14493.
XX
XX 26-JUN-1998; 98US-0105839.
XX
XX (LUDWIG) LUDWIG INST CANCER RES.
XX
XX Tureci O, Sahin U, Pfeundschnuh M, Rammensee G, Stevanovic S;
XX Chen Y, Gure A, Old LJ;
XX WPI; 2000-170933/15.
XX
XX Determining the possible presence of breast, endometrial, colorectal,
XX lung, bladder or head-neck cancer -
XX
XX Example 12; Page 21; 40pp; English.
XX
XX A method has been developed for determining the possible presence of a
XX cancer, which is not melanoma or synovial sarcoma. The method comprises
XX assaying a sample taken from the subject to determine the expression of
XX an SXX gene, and determining the expression of SXX1 gene indicates possible
XX presence of cancer. Expression of SXX1 gene indicates possible
XX presence of breast, endometrial, colorectal, lung, bladder or head-neck
XX cancer. These cancers are also detected by SXX2 and SXX4 gene
XX expression. SXX2 gene expression additionally indicates possible presence
XX of lymphoma, renal cell cancer, glioma and prostate cancer. Expression
XX of SXX4 gene also indicates possible presence of ovarian or stomach
XX cancer. SXX5 gene expression indicates the same cancers as SXX1, except
XX breast cancer. Determining expression of SXX gene can be used to monitor
XX progress of melanoma or synovial sarcoma, which is not cancer. The
XX SXX-derived peptide complex stimulates proliferation of cytolytic T
XX cells. This is useful for treating cancer, especially melanoma. AAY78464
XX to AAY74468 represent specifically claimed HLA binding peptides for use
XX in the method of the invention. AA28452 to AA28465 represent PCR
XX primers used in the isolation of SXX genes in the exemplification of the
XX present invention. AAY78469 to AAY78500, and AAY79684 to AAY7972
XX represent peptides derived from SXX proteins or NY-ESO-1, which are used
XX in the exemplification of the present invention.
XX
XX Sequence 9 AA;
XX
XX Query Match 100.0%; Score 51; DB 21; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 SLIMWITQC 9
XX |||||
XX 1 SLIMWITQC 9
XX
XX RESULT 7
XX ID AAY52432 standard; peptide; 9 AA.
XX
XX AC AAY52432;
XX
XX DT 15-FEB-2000 (first entry)
XX
XX DE Human tumour antigen NY-ESO-1 peptide #5.
XX
XX KM Cancer; tumour; antigen; MHC; major histocompatibility complex; Class I;
XX T-cell; cytotoxic; stimulation; proliferation; treatment;
XX diagnosis; prevention; melanoma; breast cancer; ovarian cancer;
XX prostate cancer; hepatoma; thyroid cancer; bladder cancer; lung cancer;
XX lymphoma.
XX
XX OS Synthetic.

```

```

OS Homo sapiens.
XX
XX MO200000824-A1.
XX
XX 28-OCT-1999.
XX
XX 24-MAR-1999; 99WO-US06675.
XX
XX 17-APR-1998; 98US-0062422.
XX
XX 02-OCT-1998; 98US-0165546.
XX
XX (LUDWIG) LUDWIG INST CANCER RES.
XX
XX Stockert E, Jager E, Chen Y, Scanlan M, Alexander K, Old LJ;
XX Gure A, Ritter G;
XX WPI; 2000-038483/03.
XX
XX Novel peptides which bind to MHC class I and MHC class II molecules,
XX useful for therapeutic and diagnostic purposes -
XX
XX Claim 60; Page 18; 49pp; English.
XX
XX Peptides #4-#7 (AAY52431-Y52434) are peptides derived from the human
XX tumour antigen, NY-ESO-1 (AAY52430) which contain the motif SLIMWIT
XX (AAY52441). These sequences can bind to MHC (major histocompatibility
XX Class I HLA-A2 molecules, thereby stimulating proliferation of
XX cytotoxic T-cells. cDNA encoding NY-ESO-1 was initially isolated
XX from an oesophagus squamous cell cancer cDNA library. Tissue
XX localisation studies revealed it to be expressed at high levels
XX in normal ovary and testis but not in normal colon, kidney, liver,
XX brain, oesophagus and skin. It was expressed in certain tumours and
XX tumour cell lines with some degree of frequency - these included
XX melanoma specimens and cell lines, and breast and bladder cancer
XX specimens, with expression in other tumour types being sporadic.
XX These NY-ESO-1-derived peptides may be used in methods and
XX compositions used for the treatment, diagnosis and prevention of
XX cancers (such as melanoma, breast cancer, prostate cancer, lung
XX cancer, hepatoma, ovarian cancer, thyroid cancer, bladder cancer,
XX or lymphoma) and to stimulate the proliferation of T cells.
XX
XX Sequence 9 AA;
XX
XX Query Match 100.0%; Score 51; DB 21; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 SLIMWITQC 9
XX |||||
XX 1 SLIMWITQC 9
XX
XX RESULT 8
XX ID AAG67166 standard; peptide; 9 AA.
XX
XX AC AAG67166;
XX
XX DT 13-NOV-2001 (first entry)
XX
XX DE Cancer testis tumour antigen NY-ESO-1 derived CTL-stimulating peptide.
XX
XX KM Cancer testis tumour antigen; NY-ESO-1; LAGE-2; human leukocyte antigen;
XX HLA; HLA binding peptide; major histocompatibility complex; MHC;
XX tumour; cancer; testis tumour.
XX
XX OS Homo sapiens.
XX
XX KM WO200162917-A1.
XX
XX PD 30-AUG-2001.
XX
XX 22-JAN-2001; 2001WO-US02126.

```

RESULT 4
 AAB22791 ID AAB22791 standard; peptide; 9 AA.
 XX
 AC AAB22791;
 XX
 DT 22-DEC-2000 (first entry)
 XX
 DE NY-ESO-1 peptide epitope, SEQ ID NO:2.
 XX
 KW NY-ESO-1; epitope; CTL response; cytotoxic T lymphocyte; vaccine;
 KW immunogenic; adjuvant coadministration; microbial infection;
 KW tuberculosis; HIV; hepatitis B virus; hepatitis C virus; cancer.
 XX
 OS Unidentified.
 XX
 PN WO200048630-A1.
 XX
 PD 24-AUG-2000.
 XX
 PF 17-FEB-2000; 2000WO-AU00110.
 XX
 PR 17-FEB-1999; 99AU-0008735.
 XX
 PR 27-JUL-1999; 99AU-0001861.
 XX
 PA (CSLC-) CSL LTD.
 XX
 PI Cox JC, Drane DP;
 XX
 DR WPI; 2000-571930/53.
 XX
 PT Immunogenic complexes comprising negatively charged organic carrier
 PT adjuvants and positively charged antigens for use as vaccines against
 PT microbial infection and cancer -
 XX
 PS Example 4; Fig 5c; 11pp; English.
 XX
 CC The invention relates to a novel immunogenic complex comprising a
 CC charged organic carrier and a charged antigen which are
 CC electrostatically associated. The complex induces a cytotoxic T
 CC lymphocyte (CTL) response. The complex and/or vaccine can be used to
 CC treat a disease in a mammal, where the complex/vaccine elicits, induces
 CC or otherwise facilitates an immune response which inhibits, halts, delays
 CC or prevents the onset or progression of the disease condition. In
 CC particular, the disease is a condition resulting from a microbial
 CC infection or cancer. Microbial infections which may be treated using the
 CC immunogenic complex include human immunodeficiency virus (HIV), hepatitis
 CC B, hepatitis C, tuberculosis or a parasitic condition, and cancers which
 CC may be treated include melanoma, prostate cancer or breast cancer. The
 CC complexes and vaccines simultaneously co-deliver antigen and adjuvant to
 CC the same antigen presenting cell, which is often essential for induction
 CC of appropriate immune responses. Sequences AAB22790-B22791 represent
 CC peptide epitopes of the positively charged protein NY-ESO-1 used in an
 CC exemplification of the invention.
 XX
 SQ Sequence 9 AA;
 XX
 Query Match 100.0%; Score 51; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 SLWMWITQC 9
 DB 1 SLWMWITQC 9
 XX
 RESULT 5
 AAB02631 ID AAB02631 standard; Peptide; 9 AA.
 XX
 AC AAB02631;
 XX

XX
 DT 18-AUG-2000 (first entry)
 XX
 DE Tumour associated peptide antigen from NY-ESO-1 #2.
 XX
 KW MAGE-A3; HLA class II; human leukocyte antigen; antibody; vaccine;
 KW cancer; human; tumour; tumour associated gene product.
 XX
 OS Homo sapiens.
 XX
 PN WO200020581-A1.
 XX
 PD 13-APR-2000.
 XX
 PF 15-SEP-1999; 99WO-US21230.
 XX
 PR 05-OCT-1998; 98US-0166448.
 XX
 PA (LUDM-) LUDMIG INST CANCER RES.
 PA (UYVR-) UNIV VIRIB BRUSSEL.
 XX
 PI Chaux P, Strobant V, Boon-Faljeur T, Van Der Bruggen P;
 PI Schultz ES, Van Snick J, Leche B, Thielemans K, Cortals J;
 PI Heitman C;
 XX
 DR WPI; 2000-317713/27.
 XX
 PT New MAGE-A3 class II binding peptides, useful to diagnose and treat
 PT tumours, are fragments of MAGE-A3 which bind to and are presented to T
 PT lymphocytes by human leukocyte antigen class II molecules -
 XX
 PS Disclosure; Page 33; 119pp; English.
 XX
 CC The present invention relates to MAGE-A3 (tumour associated gene
 CC product) human leukocyte antigen (HLA) class II-binding peptides (see
 CC AAB02566-B02595, and AAB02633-B02637). These peptides are presented to T
 CC cells in the context of HLA class II molecules. The peptides stimulate
 CC the activity and proliferation of CD4+ T lymphocytes. The invention also
 CC includes nucleotide sequences encoding MAGE-3A peptides (see AAA37928
 CC and AAA37938-A37940). The peptides and nucleotide sequences can be used
 CC to create antibodies against the MAGE-A3 peptides, the antibodies, the
 CC peptides and nucleotide sequences can be used to create a vaccine. The
 CC peptides are used to diagnose, or treat a disorder characterized by
 CC expression of MAGE-3, particularly cancer. The method can also be used
 CC in the diagnosis of disorders associated with MAGE-3 expression. Included
 CC in the invention are other human tumour antigens (see AAB02596-B02637),
 CC and PCR primers used in the course of the invention (see AAA37929-A37937
 CC and AAA37941-A37942).
 XX
 SQ Sequence 9 AA;
 XX
 Query Match 100.0%; Score 51; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 SLWMWITQC 9
 DB 1 SLWMWITQC 9
 XX
 RESULT 6
 AAY78470 ID AAY78470 standard; Peptide; 9 AA.
 XX
 AC AAY78470;
 XX
 DT 10-MAY-2000 (first entry)
 XX
 DE NY-ESO-1 derived peptide #2.
 XX
 KW Cancer; SSX family; SSX-1; SSX-2; SSX-3; SSX-4; SSX-5; NY-ESO-1;
 KW HLA binding; human leukocyte antigen; cytolytic T cell; CTL;
 KW cytostatic; melanoma; synovial sarcoma.
 XX

for the diagnosis and treatment of cancers and for monitoring cancer therapy

Claim 33, Page 17, 49pp; English.

Peptides AAW62585-87 are derived from cancer associated antigen NY-ESO-1, and are stimulators of cytotoxic T-cells. The specification describes a method for determining regression, progression of onset of a cancerous condition, comprising monitoring a sample from a patient with the cancerous condition for a parameter selected from NY-ESO-1 protein, a peptide derived from NY-ESO-1 protein and cytolytic T cells specific for the peptide and an MHC molecule with which it non-covalently complexes. Methods for the treatment of a cancerous condition are also described. The NY-ESO-1 protein and peptides derived from it can be used for diagnosis and treatment of cancers and to monitor the efficacy of a therapeutic regime.

Sequence 9 AA;

Query Match	100.0%	Score 51	DB 19	Length 9
Best Local Similarity	100.0%	Pred. No. 9.3e+05		
Matches 9	Conservative 0	Mismatches 0	Indels 0	Gaps 0
OY	1 SLWMWITOC 9			
Db	1 SLWMWITOC 9			

RESULT 2

AA01762

AA01762 standard; peptide; 9 AA.

AA01762;

25-JUN-1999 (first entry)

Exemplary antigenic peptide derived from NY-ESO-1.

MAGE-3, tumour associated gene; human leucocyte antigen Class II; autologous CD4+ cell; MAGE-3 related disease; cancer; melanoma; osteosarcoma; leukemia; carcinoma.

Homo sapiens.

WO9914326-A1.

25-MAR-1999.

04-SEP-1998; 98WO-US18601.

12-SEP-1997; 97US-0928615.

(LUDW-) LUDWIG INST CANCER RES.

(UYVR-) UNIV VRIJE BRUSSEL.

Boon-Falleur T, Chaux P, Corthals J, Heitman C; Lutten R, Stroobant V, Thielemans K, Van Der Bruggen P; WPT; 1999-244031/20.

Isolated peptides that bind to human leucocyte antigen class II molecules

Disclosure; Page 29, 89pp; English.

The present sequence represents an exemplary tumour associated peptide antigen. The specification describes a MAGE-3 tumour associated gene. Peptides (AA01761-25) that bind human leucocyte antigen (HLA) Class II molecules can be derived from the MAGE-3 protein. These peptides and autologous CD4+ cells that bind to a complex of MAGE-3 peptide and HLA Class II, are used to treat MAGE-3 related diseases, and various forms of carcinoma (e.g. melanoma, osteosarcoma, leukemia and

CC	specific antibodies. Detection of of the peptides, e.g. in binding
CC	assays, particularly with antibodies, is used for diagnosis of such
CC	diseases.
SQ	Sequence 9 AA;
QY	Query Match 100.0%; Score 51; DB 20; Length 9;
	Best Local Similarity 100.0%; Pred. No. 9.3e+05;
	Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0
DB	1 SLIMWITOC 9
DB	1 SLIMWITOC 9
DR	Result 3 AAB08703
AC	AAB08703 standard; Peptide; 9 AA.
DT	AAB08703;
DE	02-JAN-2001 (first entry)
XX	Antigenic peptide from tumour rejection antigen NY-ESO-1.
XX	EphA3; HLA class II-binding peptide; human leukocyte antigen; antigen;
XX	CD4+ T lymphocyte; tumour associated gene; vaccine.
OS	Homo sapiens.
PN	WO20050589-A1.
PD	31-AUG-2000.
PF	18-FEB-2000; 2000WC-US04326.
PR	22-FEB-1999; 99US-0121170.
PR	08-OCT-1999; 99US-0158566.
PA	(LUDW-) LUDWIG INST CANCER RES.
PI	Chiari R, Coule P, Boon-Falleur T;
DR	WPI; 2000-572089/53.
XX	Noval tyrosine kinase receptor, EphA3 human leukocyte antigen (HLA)
PT	class II binding peptide and nucleic acid encoding the receptor, useful
PT	for diagnosing and treating conditions characterized by expression of
PT	EphA3 gene -
XX	Disclosure; Page 36; 107pp; English.
PS	AAB08668-B08704 represent antigenic peptides characteristic of tumours.
XX	The peptides may be combined in vaccines with a human EphA3 HLA (human
CC	leukocyte antigen) class II-binding peptide. EphA3 antigens, when
CC	presented by an antigen presenting cell having a HLA class II molecule,
CC	effectively induce activation and proliferation of CD4+ T lymphocytes.
CC	EphA3 is a tumour associated gene. EphA3 HLA binding peptides are used
CC	for selectively enriching a population of T lymphocytes. The peptides
CC	are also used for diagnosing a disorder characterized by EphA3 or EphA3
CC	HLA binding peptide expression. The peptides are also used to treat a
CC	disorder characterized by EphA3 expression. The EphA3 binding peptides
CC	are useful in producing vaccines and antibody.
SQ	Sequence 9 AA;
QY	Query Match 100.0%; Score 51; DB 21; Length 9;
	Best Local Similarity 100.0%; Pred. No. 9.3e+05;
	Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB	1 SLIMWITOC 9
DB	1 SLIMWITOC 9

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 17:13:40 ; Search time 32.5862 Seconds
(without alignments)
43.839 Million cell updates/sec

Title: US-10-023-182-5
Perfect score: 51
Sequence: 1 SLMMITQC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	9	19	AAW62586
2	51	100.0	9	20	AAV01762
3	51	100.0	9	21	AA808703
4	51	100.0	9	21	AA822791
5	51	100.0	9	21	AA802631
6	51	100.0	9	21	AA78470
7	51	100.0	9	21	AA52432
8	51	100.0	9	22	AA67166
9	51	100.0	9	22	AAE06850

10	51	100.0	9	22	AA865299	HLA-A2 binding NY-NY-ESO-1 human leu
11	51	100.0	9	22	AAE02130	Cytolytic T cell 1
12	51	100.0	9	22	AA101537	HLA-binding pepit
13	51	100.0	9	22	AA822017	Human NY-ESO-1 CTL
14	51	100.0	9	22	AA869948	Exemplary antigen
15	51	100.0	9	23	AA831329	Human HLA-A2.1 res
16	51	100.0	9	23	AAE26808	HLA-A2 restricted
17	51	100.0	9	24	AA021430	Human NY-ESO-1 CTL
18	51	100.0	9	24	AAU64813	Human NY-ESO-1 epi
19	51	100.0	9	21	ABP74313	Human cancer anti
20	51	100.0	10	20	AAV06006	HLA-A2 binding NY-
21	51	100.0	10	22	AA885307	HLA-A2 binding NY-
22	51	100.0	10	22	AA885309	HLA-A2 binding NY-
23	51	100.0	10	22	AA885310	Cancer associated
24	51	100.0	11	19	AA862585	Exemplary antigen
25	51	100.0	11	20	AAV01761	Antigenic peptide
26	51	100.0	11	21	AA808702	NY-ESO-1 peptide e
27	51	100.0	11	21	AA827890	Tumour associated
28	51	100.0	11	21	AA802630	NY-ESO-1 derived p
29	51	100.0	11	21	AAV78469	Human tumour anti
30	51	100.0	11	21	AAV52431	Cancer testis tumo
31	51	100.0	11	21	AA677165	Human wildtype NY
32	51	100.0	11	22	AA807777	Human NY-ESO-1 ant
33	51	100.0	11	22	AA806849	HLA-A2 binding NY-
34	51	100.0	11	22	AA885298	NY-ESO-1 human leu
35	51	100.0	11	22	AA802119	Cytolytic T cell 1
36	51	100.0	11	22	AAU01536	HLA-A2 binding pep
37	51	100.0	11	22	AA882016	Human NY-ESO-1 CTL
38	51	100.0	11	22	AA869947	Exemplary antigen
39	51	100.0	11	22	AA831327	Human NY-ESO-1 CTL
40	51	100.0	11	24	AA831328	Exemplary antigen
41	51	100.0	11	24	AAU64812	Human NY-ESO-1 CTL
42	51	100.0	12	23	AB866801	Tumour antigen NY-
43	51	100.0	14	22	AA807758	Human NY ESO-1 pro
44	51	100.0	14	22	AA807754	Human HLA-DP restr
45	51	100.0	14	22	AA807765	Human HLA-DP restr

ALIGNMENTS

RESULT 1
ID AAW62586 standard; Peptide; 9 AA.
AC AAW62586;
DT 17-SEP-1998 (first entry)
DE Cancer associated antigen peptide.
KW Cancer associated antigen; NY-ESO-1; regression; progression; onset;
XX cancer; treatment; diagnosis.
OS Synthetic.
OS Homo sapiens.
PN WO9814464-A1.
XX
XX
PD 09-APR-1998.
XX
XX
PF 15-SEP-1997; 9TWO-US16335.
XX
XX
PR 03-OCT-1996; 96US-0725182.
XX
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
XX
PI Chen Y, Drifhout JW, Gure A, Jager E, Knuth A;
PI Old LJ, Scanlan M;
XX
XX
DR WPI, 1998-286417/25.
XX
XX
PT New isolated cancer associated antigen - is used to develop products

Db 1 QLSLMMWIT 9

RESULT 15
AAB31330

ID AAB31330 standard; peptide; 9 AA.

AC AAB31330;

DT 20-APR-2001 (first entry)

DE Exemplary antigen characteristic of tumours and derived from NY-ESO-1.

KM MAGE-A1; HLA; human leukocyte antigen; CD4+ T lymphocyte; cancer;

KM MAGE-A1 HLA class II-binding protein; vaccine.

OS Homo sapiens.

PN WO200078806-A1.

PD 28-DEC-2000.

PF 14-JUN-2000; 2000WO-US16287.

PR 18-JUN-1999; 99US-0336091.

PA (LUDW-) LUDWIG INST CANCER RES.

PI Van Snick J, Lethe B, Chaux P, Boon-Falleur T, Van Der Bruggen P;

DR WPI; 2001-102698/11.

PT Novel MAGE-A1 human leukocyte antigen class II peptides which bind to
PT and are presented to the class II molecules, useful for inducing immune
PT response and treating cancers characterized by expression of MAGE-A1 -

PS Disclosure; Page 32; 78pp; English.

CC AAB31302-59 represent exemplary antigens which are characteristic
CC of tumours. They can be used to enhance the immune response of vaccines
CC comprising peptides derived from human MAGE-A1 HLA (human leukocyte
CC antigen) class II-binding protein. Peptides derived from the MAGE-A1 HLA
CC binding protein stimulate the activity and proliferation of CD4+ T
CC lymphocytes. The MAGE-A1 HLA binding protein is useful as a diagnostic
CC agent for diagnosing a disorder characterized by expression of MAGE-A1.
CC The protein is used for treating a disorder characterized by expression
CC of MAGE-A1 such as cancers e.g. melanoma, squamous cell carcinomas,
CC colorectal carcinomas, osteosarcomas, and lymphocytic leukemias.
CC Peptides derived from the MAGE-A1 HLA binding protein are useful in the
CC production of anti-tumour vaccines.

SQ Sequence 9 AA:

Query Match 100.0%; Score 46; DB 22; Length 9;

Best Local Similarity 100.0%; Pred. No. 9.3e+05; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLSLMMWIT 9
Db 1 QLSLMMWIT 9

Search completed: February 10, 2004, 17:17:29
Job time : 32.5862 secs

CC having a HLA class I molecule, effectively induce the activation
 CC and proliferation of CD8+ cytotoxic T lymphocytes (CTLs). MAGE-A12
 CC is useful for treating a subject having a disorder characterised by
 CC expression of MAGE-A12. The protein microarray comprising MAGE-A12
 CC is useful for diagnosing a disorder, especially cancer, by determining
 CC the binding of an antibody, T lymphocytes or a HLA molecule isolated
 CC from the subject suspected of having the disorder characterised by the
 CC expression of MAGE-A12. MAGE-A12 is useful for treating cancers,
 CC including bladder carcinomas, melanomas, oesophageal, lung, head and
 CC neck, breast, colorectal carcinomas, myelomas, brain tumours, sarcomas,
 CC prostate and renal carcinomas and to produce antibodies, MAGE-A12
 CC antibodies are useful for diagnosing disorders characterised by
 CC expression of MAGE-A12 immunogenic polypeptide. These MAGE-A12 peptides
 CC are used as vaccines. They are also used in gene therapy.
 CC The present sequence is an antigenic peptide derived from NY-ESO-1. This
 CC peptide which is characteristic of tumours is presented by HLA-A2 MHC
 CC (major histocompatibility complex) and is recognised by CTLs.

XX Sequence 9 AA;

Query Match 100.0%; Score 46; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLSLMMWT 9
 DB 1 QLSLMMWT 9

RESULT 13

AAB82018
 ID AAB82018 standard; Peptide; 9 AA.

AC AAB82018;

DT 12-JUN-2001 (first entry)

XX HLA- binding peptide derived from NY-ESO-1.

XX Multiple myeloma; tumour rejection antigen precursor; MAGE; BAGE; GAGE;
 KM LAGE; NY-ESO-1; PRAME; DAGE; human; HLA.

XX Homo sapiens.

XX US6210886-B1.

XX 03-APR-2001.

XX 30-OCT-1998; 98US-0183931.

XX 04-FEB-1998; 98US-0018422.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Van Baren N, Brasseur F, Boon-Falleur T;

XX WPI; 2001-283628/30.

XX Detecting multiple myeloma in a patient, comprises contacting a nucleic
 PT acid containing sample taken from bone marrow or blood with a
 PT hybridization probe specific for a tumor rejection antigen precursor -

XX Example 3; Column 11; 16pp; English.

XX The present invention relates to a method for detecting multiple myeloma.
 CC The method comprises contacting a nucleic acid containing a sample taken
 CC from a bone marrow or blood of a patient, with a hybridisation probe
 CC specific for a tumour rejection antigen precursor. Tumour rejection
 CC antigen precursors used in the present invention are the MAGE family,
 CC BAGE, GAGE, LAGE, NY-ESO-1 and PRAME (previously referred to as DAGE).
 CC Expression of the tumour rejection antigen precursor indicates possible
 CC multiple myeloma in the patient. The method can also be used for
 CC monitoring the disease progress and course of therapeutic regime. The

CC present sequence is a peptide derived from a tumour rejection antigen
 CC precursor, which was used in the method of the present invention.

XX Sequence 9 AA;

Query Match 100.0%; Score 46; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLSLMMWT 9
 DB 1 QLSLMMWT 9

RESULT 14

AAB69949
 ID AAB69949 standard; Peptide; 9 AA.

AC AAB69949;

DT 27-APR-2001 (first entry)

XX Human NY-ESO-1 CTL stimulating peptide #3.

XX Human; NY-ESO-1; HLA; human leukocyte antigen; CTL; cytotoxic T cell;

XX HLA-A2; HLA-DR53; melanoma; adenocarcinoma; bladder carcinoma;
 XX non-small cell lung carcinoma; tumour status determination.

XX Homo sapiens.

XX WO200107917-A1.

XX 01-FEB-2001.

XX 14-JUL-2000; 2000WO-US19220.

XX 23-JUL-1999; 99US-0359503.

XX (LUDW-) LUDWIG INST CANCER RES.

XX (SLOK) SLOAN KETTERING INST CANCER RES.

XX (CORR) CORNELL RES FOUND INC.

XX Jager E, Stockert E, Old LJ, Knuth A, Chen Y, Scanlan M;

XX WPI; 2001-182822/18.

XX Method useful for determining the status (e.g. progression, regression
 PT or stability) of the disease) of a cancerous condition, involves
 PT determining the levels of NY-ESO-1 specific antibodies in a sample
 PT taken from a patient -

XX Example 13; Page 24; 50pp; English.

XX The present sequence is given in a specification relating to a method
 CC for determining the status of a cancerous condition in a patient
 CC with a tumour that expresses NY-ESO-1. The method comprises assaying a
 CC sample taken from the patient for antibodies that specifically bind to
 CC the NY-ESO-1 and comparing the value obtained to a prior value obtained
 CC from assay of a prior sample taken from the patient. Any difference
 CC between the values is indicative of a change in status of the cancerous
 CC condition. The method is useful for determining whether a cancerous
 CC condition is progressing, regressing or remaining stable, in particular
 CC in patients receiving treatment for a melanoma, adenocarcinoma,
 CC non-small cell lung carcinoma or bladder carcinoma.

XX Sequence 9 AA;

Query Match 100.0%; Score 46; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLSLMMWT 9
 DB 1 QLSLMMWT 9

XX Functional variants and isolated mimetics of a MAGE-A1 HLA-B35 or
PT HLA-B44 binding peptide, or of a MAGE-A3 HLA-B35 binding peptide, used
PT in diagnosis and treatment of a disorder characterized by expression of
PT MAGE-A1 or -A3 -
XX
XX Disclosure; Page 28; 103pp; English.
XX
XX The invention relates to functional variants and isolated mimetics of a
CC MAGE-A1 human leukocyte antigen (HLA)-B35 or HLA-B44 binding peptide, of
CC or of a MAGE-A3 HLA-B35 binding peptide, identified by methods described
CC in the specification. MAGE genes encode tumour rejection antigens
CC (TRAs) presented to T lymphocytes by HLA-B35 and HLA-B44 molecules.
CC The MAGE antigenic peptide acts by binding to HLA molecules
CC on tumour cells and stimulating recognition of these cells and thus
CC signalling them to the immune system for destruction. The peptide when
CC presented by HLA molecule induces the activation and stimulation of
CC CD8+ cytotoxic T lymphocytes. The MAGE antigenic peptide is used to
CC treat and diagnose disorders characterised by expression of MAGE-A1
CC or -A3. Disorders include cancers e.g melanomas, oesophageal, lung,
CC head and neck, breast, colorectal, prostate, renal, bladder,
CC hepatocellular, papillary thyroid and gastric carcinomas, myelomas,
CC brain tumours, sarcomas, seminomas, and ovarian tumours. The present
CC sequence is human NY-ESO-1 tumour associated antigenic peptide
CC presented by major histocompatibility complex (MHC) HLA-A2. The
CC antigenic peptide is used in combination with peptides of the
CC invention for inducing an immune response.
XX
XX Sequence 9 AA;
SQ
Query Match 100.0%; Score 46; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QLSLMWIT 9
DB 1 QLSLMWIT 9
RESULT 11
AAB85300
ID AAB85300 standard; peptide; 9 AA.
XX
XX AAB85300;
AC
XX
XX 17-SEP-2001 (first entry)
DT
XX
XX HLA-A2 binding NY-ESO-1 peptide #3.
DE
XX
XX NY-ESO-1; human leukocyte antigen; HLA; lysis; cytolytic T cell; CTL;
KW HLA-A2; T-cell sorter; tumor; immune tetramer.
XX
XX Homo sapiens.
OS
XX
XX WO200136453-A2.
PN
XX
XX 25-MAY-2001.
PD
XX
XX 08-NOV-2000; 2000WO-US42010.
PF
XX
XX 15-NOV-1999; 9US-0440621.
PR
XX
XX 25-FEB-2000; 2000US-0514036.
PR
XX
XX 29-SEP-2000; 2000US-0676005.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
PA
XX
XX (UYOX-) UNIV OXFORD.
PA
XX
XX Valmori D, Cerottini J, Romero P, Cerundolo V;
PI
XX
XX WPI; 2001-451454/48.
DR
XX
XX Novel isolated NY-ESO-1 nonapeptide useful for determining if a cell
PT presents human leukocyte antigen-A2 molecule on its surface, binds to

PT human leukocyte antigen molecules and provokes lysis by cytolytic T
PT cells -
XX
XX Example 1; Page 4; 38pp; English.
XX
XX The invention provides NY-ESO-1 peptide derivatives which bind to human
CC leukocyte antigen (HLA) molecules and provokes lysis by cytolytic T cells
CC (CTLs). The NY-ESO-1 nonapeptide is of formula SLMWITX, where X is an
CC amino acid having an uncharged polar side chain. The NY-ESO-1 peptide
CC derivatives are useful for determining if a cell presents an HLA-A2
CC molecule on its surface, by contacting a sample containing the cell with
CC the peptide or its derivative, and determining binding between them,
CC where the binding is indicative of HLA-A2 on the surface of the cell.
CC The NY-ESO-1 peptides and analogues are useful therapeutically, for
CC administration to a patient who is HLA-A2 positive and expresses NY-ESO-1
CC in connection with the pathology, as well as diagnostically, i.e. to
CC determine if HLA-A2 positive cells are present, or if relevant CTLs are
CC present. They are also useful for determining the presence of CTLs in a
CC sample. The peptides are useful as T-cell sorters, when incorporated into
CC immune tetramers. The present sequence represents a NY-ESO-1 peptide that
CC can bind to HLA-A2 molecule.
XX
XX Sequence 9 AA;
SQ
Query Match 100.0%; Score 46; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QLSLMWIT 9
DB 1 QLSLMWIT 9
RESULT 12
AAE02121
ID AAE02121 standard; peptide; 9 AA.
XX
XX AAE02121;
AC
XX
XX 31-JUL-2001 (first entry)
DT
XX
XX NY-ESO-1 human leukocyte antigen-A2-binding peptide #3.
DE
XX
XX Human; cytostatic; immunogen; NY-ESO-1; human leukocyte antigen;
KW HLA; CD8; cytotoxic T lymphocyte; cancer; carcinoma; melanoma; myeloma;
KW brain tumour; sarcoma; vaccine; gene therapy.
XX
XX Homo sapiens.
OS
XX
XX WO200129220-A2.
PN
XX
XX 26-APR-2001.
PD
XX
XX 19-OCT-2000; 2000WO-US28852.
PF
XX
XX 19-OCT-1999; 9US-0160374.
PR
XX
XX 01-FEB-2000; 2000US-0179570.
PR
XX
XX (LUDW-) LUDWIG INST CANCER RES.
PA
XX
XX Heidecker L, Van Den Eynde B, Boon-Falleur T, Brasseur F;
PI
XX
XX WPI; 2001-328498/34.
DR
XX
XX New antigenic peptides derived from MAGE-A12 polypeptides, useful for
PT diagnosis and treatment of cancer, such as bladder, lung, breast,
PT brain, prostate and renal carcinomas -
PT
XX
XX Disclosure; Page 21; 69pp; English.
XX
XX The patent discloses antigenic peptides derived from MAGE-A12
CC protein and presented by human leukocyte antigens (HLAs). These
CC antigenic peptides when presented by an antigen presenting cell

OS Homo sapiens.
 XX WO200170772-A2.
 XX 27-SEP-2001.
 PD
 XX 22-MAR-2001; 2001WO-FR00872.
 XX
 XX 23-MAR-2000; 2000FR-0003711.
 XX
 XX (FABR) FABRE MEDICAMENT SA PIERRE.
 XX
 XX Klinguer-Hamou C, Corvaia N, Beck A, Goetsch L,
 XX WPI; 2001-611470/70.
 DR
 XX Stabilized pharmaceutical containing N-terminal glutamic acid or
 PT glutamine, useful e.g. in anti-melanoma vaccines, is an addition salt
 PT with strong acid
 XX
 XX Claim 9, Page 11; 149pp, French.
 PS
 XX The present invention describes a pharmaceutical compound (I) that
 CC contains an N-terminal glutamic acid (Glu) or glutamine (Gln) residue
 CC in the form of an addition salt with a strong, physiologically
 CC acceptable acid (II). Also described are: (a) a pharmaceutical
 CC composition containing at least one (I); (b) a vaccine containing at
 CC least one (I) where this is a major histocompatibility complex (MHC)
 CC ligand (Ia); (c) a method for in vitro diagnosis of diseases associated
 CC with the presence of (Ia); (d) a kit for method (c) that includes a (Ia);
 CC and (e) a process for preparing (I). (I) has immunomodulator, endocrine,
 CC antiallergic, neuroprotectant, virucidal, bactericidal, antiparasitic,
 CC fungicidal and cytostatic activities. (I) are useful in human or
 CC veterinary medicine, in pharmaceutical compositions (for treating immune
 CC disorders, e.g. immune deficiency, autoimmune states, hypersensitivity,
 CC allergy, graft rejection, infection, hormonal disorders and central
 CC nervous system diseases), also, where (I) is a MHC ligand (Ia), in
 CC vaccines for treatment or prevention of: (i) viral, bacterial, parasitic
 CC or fungal infections; or (ii) of cancers. A particular application is in
 CC anti-melanoma vaccines. (I) are also useful for in vitro diagnosis of
 CC diseases associated with interactions between MHC and (I), e.g. melanoma
 CC and human immunodeficiency virus infection. AAY98898 to AAY95952
 CC represent peptides which can be used in pharmaceutical compounds from
 CC the present invention.
 CC
 CC Sequence 9 AA;
 SQ
 Query Match 100.0%; Score 46; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QLSILMWIT 9
 DB 1 QLSILMWIT 9

RESULT 9
 AAG67167
 ID AAG67167 standard; peptide; 9 AA.
 AC
 XX AAG67167;
 AC
 DT 13-NOV-2001 (first entry)
 XX
 XX Cancer testis tumour antigen NY-ESO-1 derived CTL-stimulating peptide.
 DE
 XX Cancer testis tumour antigen; NY-ESO-1; LAGE-2; human leukocyte antigen;
 KW HLA; HLA binding peptide; major histocompatibility complex; MHC;
 KW tumour; cancer; testis tumour.
 XX
 XX Homo sapiens.
 OS
 XX WO200162917-A1.
 PN

XX 30-AUG-2001.
 PD
 XX 22-JAN-2001; 2001WO-US02126.
 XX
 XX 22-FEB-2000; 2000US-0510635.
 XX
 XX (LUDW-) LUDWIG INST CANCER RES.
 XX
 XX Leche B, Boon-Falleur T;
 XX WPI; 2001-550091/61.
 DR
 XX Genomic sequences of tumour associated antigen EY-ESO-1 (LAGE-2) useful
 PT for diagnosing testicular tumours -
 PT
 XX Example 12; Page 24; 50pp; English.
 PS
 XX The present sequence represents a peptide which is derived from cancer
 CC testis tumour antigen NY-ESO-1 (also called LAGE-2). The peptide that
 CC stimulates cytolytic T cell lines (CTLs). NY-ESO-1 is a molecule that
 CC is processed to at least one human leukocyte antigen (HLA) binding
 CC peptide, which binds to Class I and Class II major histocompatibility
 CC complex (MHC). NY-ESO-1 is expressed in tumour mRNA and in testis,
 CC but not normal colon, kidney, liver or brain tissue. The presence or
 CC level of expression of NY-ESO-1 may be assayed for the diagnosis of
 CC cancer, especially testis tumours.
 CC
 CC Sequence 9 AA;
 SQ
 Query Match 100.0%; Score 46; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QLSILMWIT 9
 DB 1 QLSILMWIT 9

RESULT 10
 AAE06851
 ID AAE06851 standard; peptide; 9 AA.
 AC
 XX AAE06851;
 AC
 DT 16-OCT-2001 (first entry)
 XX
 XX Human NY-ESO-1 antigenic peptide #3.
 DE
 XX MAGE antigenic peptide; Human leukocyte antigen; HLA-B35; HLA-B44;
 KW tumour cell; immunostimulant; antigen presentation; cancer; melanoma;
 KW CD8+ cytotoxic T lymphocyte; colorectal; prostate; gastric carcinoma;
 KW myeloma; brain tumour; sarcoma; seminoma; ovarian tumour; cytostatic;
 KW gene therapy; tumour rejection antigen; TRA; human; NY-ESO-1;
 KW MHC; major histocompatibility complex.
 XX
 XX Homo sapiens.
 OS
 XX WO200153833-A1.
 PN
 XX 26-JUL-2001.
 PD
 XX 19-JAN-2001; 2001WO-US02008.
 XX
 XX 20-JAN-2000; 2000US-0177242.
 XX
 XX 25-OCT-2000; 2000US-0243212.
 XX
 XX (LUDW-) LUDWIG INST CANCER RES.
 XX
 XX Luiten R, Boon-Falleur T, Van Der Bruggen P, Stroobant V,
 PI Demotte N, Schultz E;
 PI WPI; 2001-488724/53.
 DR

KM Cancer; SSX family; SSX-1; SSX-2; SSX-3; SSX-4; SSX-5; NY-ESO-1;
 KM HLA binding; human leukocyte antigen; cytolytic T cell; CTL;
 KM Cytostatic; melanoma; synovial sarcoma.
 OS Homo sapiens.
 XX
 XX WO200000824-A1.
 XX
 XX PD 06-JAN-2000.
 XX
 XX PF 25-JUN-1999; 99WO-US14493.
 XX
 XX PR 26-JUN-1998; 98US-0105839.
 XX
 XX (LUDW-) LUDWIG INST CANCER RES.
 PA
 PI Thureci O, Sahin U, Pfeundschn M, Rammensee G, Stevanovic S;
 PI Chen Y, Gure A, Old LJ;
 XX
 XX WPI; 2000-170933/15.
 DR
 XX
 PT Determining the possible presence of breast, endometrial, colorectal,
 PT lung, bladder or head-neck cancer -
 XX
 XX Example 13; Page 26; 40pp; English.
 PS
 XX A method has been developed for determining the possible presence of a
 CC cancer, which is not melanoma or synovial sarcoma. The method comprises
 CC assaying a sample taken from the subject to determine the expression of
 CC an SSX gene, and determining the expression as a determination of the
 CC possible presence of cancer. Expression of SSX1 gene indicates possible
 CC presence of breast, endometrial, colorectal, lung, bladder or head-neck
 CC cancer. These cancers are also detected by SSX2 and SSX4 gene
 CC expression. SSX2 gene expression additionally indicates possible presence
 CC of lymphoma, renal cell cancer, glioma and prostate cancer. Expression
 CC of SSX4 gene also indicates possible presence of ovarian or stomach
 CC cancer. SSX5 gene expression indicates the same cancers as SSX1, except
 CC breast cancer. Determining expression of SSX gene can be used to monitor
 CC progress of melanoma or synovial sarcoma, which is not cancer. The
 CC SSX-derived peptide complex stimulates proliferation of cytolytic T
 CC cells. This is useful for treating cancer, especially melanoma. AA78464
 CC to AA78468 represent specifically claimed HLA binding peptides for use
 CC in the method of the invention. AA288452 to AA288465 represent PCR
 CC primers used in the isolation of SSX genes in the exemplification of the
 CC present invention. AA78469 to AA78500, and AA79684 to AA79762
 CC represent peptides derived from SSX proteins or NY-ESO-1, which are used
 CC in the exemplification of the present invention.
 CC
 XX
 SQ Sequence 9 AA;

Query Match 100.0%; Score 46; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLSILMWIT 9
 Db 1 QLSILMWIT 9

RESULT 7
 AA52433
 ID AAY52433 standard; peptide; 9 AA.
 AC
 XX AAY52433;
 XX
 DT 15-FEB-2000 (first entry)
 XX
 DE Human tumour antigen NY-ESO-1 peptide #6.
 XX
 XX Cancer; tumour; antigen; MHC; major histocompatibility complex; Class I;
 KM T-cell; cytotoxic; stimulation; proliferation; treatment;
 KM diagnosis; prevention; melanoma; breast cancer; ovarian cancer;
 KM prostate cancer; hepatoma; thyroid cancer; bladder cancer; lung cancer;

KM lymphoma.
 XX
 XX Synthetic.
 OS Homo sapiens.
 XX
 XX WO9853938-A1.
 XX
 XX PD 28-OCT-1999.
 XX
 XX PF 24-MAR-1999; 99WO-US06875.
 XX
 XX PR 17-APR-1998; 98US-0062422.
 XX
 XX PR 02-OCT-1998; 98US-0165546.
 XX
 XX (LUDW-) LUDWIG INST CANCER RES.
 PA
 PI Stockert E, Jager E, Chen Y, Scanlan M, Alexander K, Old LJ;
 PI Gure A, Ritter G;
 XX
 XX WPI; 2000-036483/03.
 DR
 XX
 PT Novel peptides which bind to MHC class I and MHC class II molecules,
 PT useful for therapeutic and diagnostic purposes -
 XX
 XX Claim 60; Page 18; 49pp; English.
 PS
 XX

Peptides #4-#7 (AAY52431-V52434) are peptides derived from the human
 CC tumour antigen, NY-ESO-1 (AAY52430) which contain the motif LLMWIT
 CC (AAY52441). These sequences can bind to MHC (major histocompatibility
 CC class I HLA-A2 molecules, thereby stimulating proliferation of
 CC cytotoxic T-cells. cDNA encoding NY-ESO-1 was initially isolated
 CC from an oesophagus squamous cell cancer cDNA library. Tissue
 CC localization studies revealed it to be expressed at high levels
 CC in normal ovary and testis but not in normal colon, kidney, liver,
 CC brain, oesophagus and skin. It was expressed in certain tumours and
 CC tumour cell lines with some degree of frequency - these included
 CC melanoma specimens and cell lines, and breast and bladder cancer
 CC specimens, with expression in other tumour types being sporadic.
 CC These NY-ESO-1-derived peptides may be used in methods and
 CC compositions used for the treatment, diagnosis and prevention of
 CC cancers (such as melanoma, breast cancer, prostate cancer, lung
 CC cancer, hepatoma, ovarian cancer, thyroid cancer, bladder cancer,
 CC or lymphoma) and to stimulate the proliferation of T cells.

Query Match 100.0%; Score 46; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLSILMWIT 9
 Db 1 QLSILMWIT 9

RESULT 8
 AAM99363
 ID AAM99363 standard; Peptide; 9 AA.
 AC
 XX AAM99363;
 XX
 DT 07-DEC-2001 (first entry)
 XX
 DE Vaccine related MHC ligand peptide SEQ ID NO:466.
 XX
 XX Glutamic acid; glutamine; vaccine; major histocompatibility complex;
 KM MHC; immunomodulator; antitumor; endocrine; neuroprotectant;
 KM vitricidal; bactericidal; antiparasitic; fungicidal; cytostatic;
 KM medicinal; pharmaceutical; immune disorder; immune deficiency;
 KM autoimmune; hypersensitivity; allergy; graft rejection; infection;
 KM hormonal disorder; central nervous system disease; cancer; melanoma;
 KM anti-melanoma vaccine; human immunodeficiency virus.

RESULT 4
AAB02632
ID AAB02632 standard; Peptide; 9 AA.
AC AAB02632;
XX
DT 18-AUG-2000 (first entry)
XX
DE Tumour associated peptide antigen from NY-ESO-1 #3.
XX
KW MAGE-A3; HLA class II; human leukocyte antigen; antibody; vaccine;
KM cancer; human; tumour; tumour associated gene product.
OS Homo sapiens.
PN WO200020581-A1.
PD 13-APR-2000.
PF 15-SEP-1999; 99WO-US21230.
PR 05-OCT-1996; 98US-0166448.
PS (LUDW-) LUDWIG INST CANCER RES.
PA (UYVR-) UNIV WRIJE BRUSSEL.
PX Chaux P., Strobant V., Boon-Falleur T., Van Der Bruggen P.;
PI Schultz ES, Van Snick J, Letne B, Thielemans K, Cortals J,
PL Heitman C;
DR WPI; 2000-317713/27.
XX
PT New MAGE-A3 class II binding peptides, useful to diagnose and treat
PT tumour, are fragments of MAGE-A3 which bind to and are presented to T
PT lymphocytes by human leukocyte antigen class II molecules -
FS Disclosure; Page 33; 119pp; English.
XX
CC The present invention relates to MAGE-A3 (tumour associated gene
CC product) human leukocyte antigen (HLA) class II-binding peptides (see
CC AAB02586-802595, and AAB02632-802637). These peptides are presented to T
CC cells in the context of HLA class II molecules. The peptides stimulate
CC the activity and proliferation of CD4+ T lymphocytes. The invention also
CC includes nucleotide sequences encoding MAGE-3a peptides (see AA37928
CC and AA37938-A37940). The peptides and nucleotide sequences can be used
CC to create antibodies against the MAGE-A3 peptides, the antibodies,
CC peptides and nucleotide sequences can be used to create a vaccine. The
CC peptides are used to diagnose or treat a disorder characterized by
CC expression of MAGE-3, particularly cancer. The methods can also be used
CC in the diagnosis of disorders associated with MAGE-3 expression. Included
CC in the invention are other human tumour antigens (see AAB02586-802637),
CC and PCR primers used in the course of the invention (see AA37929-A37937
CC and AA37941-A37942).
XX
SQ Sequence 9 AA;
XX
Query Match 100.0%; Score 46; DB 21; Length 9;
Best Local Similarity 100.0%; Pred.No. 9.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY 1 QLSLMWT 9
DB 1 QLSLMWT 9
XX
NC AAAT78471;
XX
RESULT 5
AAAT78471
ID AAAT78471 standard; Peptide; 9 AA.
XX
NC AAAT78471;
XX

```

DT 10-MAY-2000 (first entry)
DE NY-ESO-1 derived peptide #3.
KM Cancer; SSX family; SSX-1; SSX-2; SSX-3; SSX-4; SSX-5; NY-ESO-1;
KW HLA binding; human leukocyte antigen; cytolytic T cell; CTL;
XX cytotoxic; melanoma; synovial sarcoma.
OS
XX Homo sapiens.
XX
XX WC0200000824-A1.
XX
XX 06-JAN-2000.
XX
XX 25-JUN-1999; 99WO-US14493.
XX
XX 26-JUN-1998; 98US-0105839.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Tureci O, Sahin U, Pfeundschnuh M, Rammensee G, Stevanovic S;
XX Chen Y, Gure A, Old LJ;
XX
XX WPI; 2000-170933/15.
XX
XX
XX Determining the possible presence of breast, endometrial, colorectal,
XX lung, bladder or head-neck cancer -
XX
XX Example 12; Page 21; 40pp; English.
XX
XX A method has been developed for determining the possible presence of a
XX cancer, which is not melanoma or synovial sarcoma. The method comprises
XX assaying a sample taken from the subject to determine the expression of
XX an SSX gene, and determining the expression as a determination of the
XX possible presence of cancer. Expression of SSX1 gene indicates possible
XX presence of breast, endometrial, colorectal, lung, bladder or head-neck
XX cancer. These cancers are also detected by SSX2 and SSX4 gene
XX expression. SSX2 gene expression additionally indicates possible presence
XX of lymphoma, renal cell cancer, glioma and prostate cancer. Expression
XX of SSX4 gene also indicates possible presence of ovarian or stomach
XX cancer. SSX5 gene expression indicates the same cancers as SSX1, except
XX breast cancer. Determining expression of SSX gene can be used to monitor
XX progress of melanoma or synovial sarcoma, which is not cancer. The
XX SSX-derived peptide complex stimulates proliferation of cytolytic T
XX cells. This is useful for treating cancer, especially melanoma. AA78464
XX to AA78468 represent specifically claimed HLA binding peptides for use
XX in the method of the invention. AA288452 to AA288465 represent PCR
XX primers used in the isolation of SSX genes in the exemplification of the
XX present invention. AA784459 to AA784500, and AA784564 to AA784567
XX represent peptides derived from SSX proteins or NY-ESO-1, which are used
XX in the exemplification of the present invention.
XX
XX
XX Sequence 9 AA;
XX
XX
XX Query Match 100.0%; Score 46; DB 21; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 QSLSLMWIT 9
XX 1 QSLSLMWIT 9
XX
XX
XX RESULT 6
XX ID AA79755
XX AC AA79755;
XX
XX 10-MAY-2000 (first entry)
XX
XX NY-ESO-1 derived peptide #11.
XX

```

PT for the diagnosis and treatment of cancers and for monitoring cancer
 therapy
 XX
 XX Claim 33; Page 17; 49pp; English.
 XX
 CC Peptides AAW62585-87 are derived from cancer associated antigen
 CC NY-ESO-1, and are stimulators of cytotoxic T-cells. The specification
 CC describes a method for determining regression, progression of onset of a
 CC cancerous condition, comprising monitoring a sample from a patient
 CC with the cancerous condition for a parameter selected from NY-ESO-1
 CC protein, a peptide derived from NY-ESO-1 protein and cyclolytic
 CC T cells specific for the peptide and an MHC molecule with which it
 CC non-covalently complexes. Methods for the treatment of a cancerous
 CC condition are also described. The NY-ESO-1 protein and peptides derived
 CC from it can be used for diagnosis and treatment of cancers and to
 CC monitor the efficacy of a therapeutic regime.
 XX
 SQ Sequence 9 AA;
 Query Match 100.0%; Score 46; DB 19; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 QLSLMMWIT 9
 1 |||||
 1 QLSLMMWIT 9
 DB
 RESULT 2
 AAY01763
 ID AAY01763 standard; Peptide; 9 AA.
 XX
 AC AAY01763;
 XX
 DT 25-JUN-1999 (first entry)
 XX
 DE Exemplary antigenic peptide derived from NY-ESO-1.
 XX
 KM MAGF-3; tumour associated gene; human leucocyte antigen Class II;
 KM autologous CD4+ cell; MAGF-3 related disease; cancer; melanoma;
 KM osteosarcoma; leukemia; carcinoma.
 XX
 OS Homo sapiens.
 XX
 PN WO9914326-A1.
 XX
 PD 25-MAR-1999.
 XX
 PF 04-SEP-1998; 98WO-US18601.
 XX
 PR 12-SEP-1997; 97US-0928615.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 PA (UYVR-) UNIV VRIJE BRUSSEL.
 XX
 PI Boon-Falleur T, Chaux P, Corthals J, Heitman C;
 PI Lutten R, Stroobant V, Thielemans K, Van Der Bruggen P;
 DR WPI; 1999-244031/20.
 XX
 PT Isolated peptides that bind to human leucocyte antigen class II
 PT molecules
 XX
 PS Disclosure; Page 29; 86pp; English.
 XX
 CC The present sequence represents an exemplary tumour associated peptide
 CC antigen. The specification describes a MAGF-3 tumour associated gene.
 CC Peptides (AAY01721-25) that bind human leucocyte antigen (HLA) Class II
 CC molecules can be derived from the MAGF-3 protein. These peptides and
 CC autologous CD4+ cells that bind to a complex of MAGF-3 peptide
 CC and HLA Class II, are used to treat MAGF-3 related diseases,
 CC particularly cancers (e.g. melanoma, osteosarcoma, leukemia and
 CC various forms of carcinoma). The peptides are also used to produce

CC specific antibodies. Detection of of the peptides, e.g. in binding
 CC assays, particularly with antibodies, is used for diagnosis of such
 CC diseases.
 XX
 SQ Sequence 9 AA;
 Query Match 100.0%; Score 46; DB 20; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 QLSLMMWIT 9
 1 |||||
 1 QLSLMMWIT 9
 DB
 RESULT 3
 AAB08704
 ID AAB08704 standard; Peptide; 9 AA.
 XX
 AC AAB08704;
 XX
 DT 02-JAN-2001 (first entry)
 XX
 DE Antigenic peptide from tumour rejection antigen NY-ESO-1.
 XX
 KM BphA3; HLA class II-binding peptide; human leucocyte antigen; antigen;
 KM CD4+ T lymphocyte; tumour associated gene; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO200050589-A1.
 XX
 PD 31-AUG-2000.
 XX
 PF 18-FEB-2000; 2000WO-US04326.
 XX
 PR 22-FEB-1999; 99US-0121170.
 PR 08-OCT-1999; 99US-0158566.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Chiari R, Coulie P, Boon-Falleur T;
 PI WPI; 2000-572089/53.
 XX
 DR Novel tyrosine kinase receptor, BphA3 human leucocyte antigen (HLA)
 PT class II binding peptide and nucleic acid encoding the receptor; useful
 PT for diagnosing and treating conditions characterized by expression of
 PT BphA3 gene
 XX
 PS Disclosure; Page 36; 107pp; English.
 XX
 CC AAB08668-B08704 represent antigenic peptides characteristic of tumours.
 CC The peptides may be combined in vaccines with a human BphA3 HLA (human
 CC leucocyte antigen) class II-binding peptide. BphA3 antigens, when
 CC presented by an antigen presenting cell having a HLA class II molecule,
 CC effectively induce activation and proliferation of CD4+ T lymphocytes.
 CC BphA3 is a tumour associated gene. BphA3 HLA binding peptides are used
 CC for selectively enriching a population of T lymphocytes. The peptides
 CC are also used for diagnosing a disorder characterized by BphA3 or BphA3
 CC HLA binding peptide expression. The peptides are also used to treat a
 CC disorder characterized by BphA3 expression. The BphA3 binding peptides
 CC are useful in producing vaccines and antibody.
 XX
 SQ Sequence 9 AA;
 Query Match 100.0%; Score 46; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 QLSLMMWIT 9
 1 |||||
 1 QLSLMMWIT 9
 DB

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 17:13:40 ; Search time 32.5862 Seconds
(without alignments)
43.839 Million cell updates/sec

Title: US-10-023-182-6
Perfect score: 46
Sequence: 1 QSLMLWIT 9

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	9	19	AAW62587
2	46	100.0	9	20	AAV01763
3	46	100.0	9	21	AAH08704
4	46	100.0	9	21	AAH02632
5	46	100.0	9	21	AAV78471
6	46	100.0	9	21	AAV79755
7	46	100.0	9	21	AAV52433
8	46	100.0	9	22	AAW93363
9	46	100.0	9	22	AAW67167

Result No.	Score	Query Match	Length	DB ID	Description
10	46	100.0	9	22	AAH06851
11	46	100.0	9	22	AAH85300
12	46	100.0	9	22	AAH02121
13	46	100.0	9	22	AAH82018
14	46	100.0	9	22	AAH69949
15	46	100.0	9	22	AAH31330
16	46	100.0	9	23	AAH26810
17	46	100.0	9	24	AAH04814
18	46	100.0	10	20	AAH06009
19	46	100.0	10	22	AAH67204
20	46	100.0	10	22	AAH69936
21	46	100.0	10	24	AAH64852
22	46	100.0	11	22	AAH31327
23	46	100.0	13	22	AAH07776
24	46	100.0	14	22	AAH07764
25	46	100.0	14	22	AAH07765
26	46	100.0	14	22	AAH07773
27	46	100.0	14	22	AAH07774
28	46	100.0	14	22	AAH07775
29	46	100.0	20	22	AAH07745
30	46	100.0	27	22	AAH07757
31	46	100.0	28	24	ABP74370
32	46	100.0	28	24	ABP74371
33	46	100.0	30	23	AAH85111
34	46	100.0	30	23	AAH85112
35	46	100.0	30	23	AAH85127
36	46	100.0	30	23	AAH85128
37	46	100.0	81	24	ABP72587
38	46	100.0	81	24	ABP72588
39	46	100.0	135	24	ABR48211
40	46	100.0	135	24	ABU65509
41	46	100.0	135	24	ABU65695
42	46	100.0	180	19	AAH69664
43	46	100.0	180	19	AAH69665
44	46	100.0	180	19	AAH62584
45	46	100.0	180	20	AAH05965

ALIGNMENTS

RESULT 1
AAW62587 standard; Peptide; 9 AA.
AC AAW62587;
XX
XX
DT 17-SEP-1998 (first entry)
XX
DE Cancer associated antigen peptide.
XX
XX Cancer associated antigen; NY-ESO-1; regression; progression; onset;
KW cancer; treatment; diagnosis.
XX
XX
OS Synthetic.
OS Homo sapiens.
XX
XX W09814464-A1.
XX
XX
PD 09-APR-1998.
XX
XX 15-SEP-1997; 97WO-US16335.
PF
XX
PR 03-OCT-1996; 96US-0725182.
XX
XX
XX (LUDW-) LUDWIG INST CANCER RES.
PA
XX Chen Y, Drifhout JW, Gure A, Jager E, Knuth A;
PI Old LJ, Scanlan M;
XX
XX WPI; 1998-286417/25.
DR
XX
XX
PT New isolated cancer associated antigen - is used to develop products

```

RESULT 15
AAB69947
ID AAB69947 standard; Peptide; 11 AA.
XX
AC AAB69947;
XX
DT 27-APR-2001 (first entry)
XX
DE Human NY-ESO-1 CTL stimulating peptide #1.
XX
KM Human; NY-ESO-1; HLA; human leukocyte antigen; CTL; cytotoxic T cell;
KW HLA-A2; HLA-DR5; melanoma; adenocarcinoma; bladder carcinoma;
KW non-small cell lung carcinoma; tumour status determination.
XX
OS Homo sapiens.
XX
PN WO200107917-A1.
XX
PD 01-FEB-2001.
XX
PF 14-JUL-2000; 2000WO-US19220.
XX
PR 23-JUL-1999; 99US-0359503.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
PA (SLOK) SLOAN KETERING INST CANCER RES.
PA (CORR) CORNELL RES FOUND INC.
XX
PI Jager E, Stockert E, Old LJ, Knuth A, Chen Y, Scanlan M;
XX
DR WPI; 2001-182822/18.
XX
PT Method useful for determining the status (e.g. progression, regression
PT or stability of the disease) of a cancerous condition, involves
PT determining the levels of NY-ESO-1 specific antibodies in a sample
PT taken from a patient.
XX
PS Example 13; Page 24; 50pp; English.
XX
CC The present sequence is given in a specification relating to a method
CC for determining the status of a cancerous condition in a patient
CC with a tumour that expresses NY-ESO-1. The method comprises assaying a
CC sample taken from the patient for antibodies that specifically bind to
CC the NY-ESO-1 and comparing the value obtained to a prior value obtained
CC from assay of a prior sample taken from the patient. Any difference
CC between the values is indicative of a change in status of the cancerous
CC condition. The method is useful for determining whether a cancerous
CC condition is progressing, regressing or remaining stable, in particular
CC in patients receiving treatment for a melanoma, adenocarcinoma,
CC non-small cell lung carcinoma or bladder carcinoma.
XX
SQ Sequence 11 AA:
XX
Query Match 100.0%; Score 61; DB 22; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00066;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SLMMITQCFL 11
DB 1 SLMMITQCFL 11

```

Search completed: February 10, 2004, 17:17:29
 Job time : 40.8276 secs

CC neck, breast, colorectal carcinomas, myelomas, brain tumours, sarcomas,
 CC prostate and renal carcinomas and to produce antibodies. MAGE-A12
 CC antibodies are useful for diagnosing disorders characterised by
 CC expression of MAGE-A12 immunogenic polypeptide. These MAGE-A12 peptides
 CC are used as vaccines. They are also used in gene therapy.
 CC The present sequence is an antigenic peptide derived from NY-ESO-1. This
 CC peptide which is characteristic of tumours is presented by HLA-A2 MHC
 CC (major histocompatibility complex) and is recognised by CTLs.
 SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 22; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.00066;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLIMWITQCFL 11
 |||||
 1 SLIMWITQCFL 11

RESULT 13
 AAU01536
 ID AAU01536 standard; Peptide; 11 AA.
 XX AAU01536;
 AC

DT 18-JUL-2001 (first entry)

DE Cytolytic T cell line stimulator peptide #1.

XX NY-ESO-1; human; tumour rejection antigen precursor; SSX-2; MHC Class II;
 KW major histocompatibility complex; helper T cell; HLA-DR; cancer;
 KM human leukocyte antigen-determining region; disease progression;
 KM disease regression; disease onset; body tissue; body fluid; enzyme label;
 KM radioactive label; monoclonal antibody; cytolytic T cell line.

XX Homo sapiens.

XX WC000123560-A2.

XX 05-APR-2001.

XX 26-SEP-2000; 2000WO-US26411.

XX 29-SEP-1999; 99US-0408036.

XX (LUDW-) LUDWIG INST CANCER RES.

PI Tureci O, Sahin U, Pfeundschn M;

XX WPI; 2001-266156/27.

PT Polypeptides binding to major histocompatibility complex class II human
 PT leukocyte antigen-determining region molecule having amino acid
 PT sequence found in tumour rejection antigen precursor used for
 PT stimulating proliferation of helper T cells -

XX Example 12; Page 17; 62pp; English.

XX The sequence represents a human NY-ESO-1 tumour rejection antigen
 CC precursor fragment which efficiently stimulates cytolytic T cell lines.
 CC NY-ESO-1 and SSX-2 polypeptides, or fragments of, bind to major
 CC histocompatibility complex (MHC) Class II molecules such as human
 CC leukocyte antigen-determining region (HLA-DR) molecules and stimulate
 CC proliferation of helper T cells. The peptides can be administered to an
 CC HLA-DR positive subject in order to stimulate the helper T cells. An MHC
 CC Class II HLA-DR-NY-ESO-1/SSX-2 complex expressed on the surface of a cell
 CC or present in free form is useful for this stimulation. The nucleic acid
 CC is useful for screening for a cancerous condition, which involves
 CC contacting a subject sample to a cell line transfected with the
 CC immunoreactive cell (helper T cell), where interaction is indicative of
 CC cancer. In addition, a sample from a patient (for example, a body fluid
 CC or tissue) can be monitored for the amount of the complex present in the

CC bloodstream. This is useful for determining regression, progression or
 CC onset of a cancerous condition. The method involves contacting the sample
 CC with a radioactive labelled or enzyme labelled monoclonal antibody which
 CC specifically binds with the complex.
 XX

SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 22; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.00066;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLIMWITQCFL 11
 |||||
 1 SLIMWITQCFL 11

RESULT 14
 AAB82016
 ID AAB82016 standard; Peptide; 11 AA.
 XX AAB82016;
 AC

DT 12-JUN-2001 (first entry)

DE HLA-A2 binding peptide derived from NY-ESO-1.

XX Multiple myeloma; tumour rejection antigen precursor; MAGE; BAGE; GAGE;
 KW LAGE; NY-ESO-1; PRAME; DAGE; human; HLA.

XX Homo sapiens.

XX US6210886-B1.

XX 03-APR-2001.

XX 30-OCT-1998; 98US-0183931.

XX 04-FEB-1998; 98US-0018422.

XX (LUDW-) LUDWIG INST CANCER RES.

PI Van Baren N, Brasseur F, Boon-Falleur T;

XX WPI; 2001-289628/30.

PT Detecting multiple myeloma in a patient, comprises contacting a nucleic
 PT acid containing sample taken from bone marrow or blood with a
 PT hybridization probe specific for a tumor rejection antigen precursor -
 XX

XX Example 3; Column 11; 16pp; English.

XX The present invention relates to a method for detecting multiple myeloma.
 CC The method comprises contacting a nucleic acid containing a sample taken
 CC from a bone marrow or blood of a patient, with a hybridisation probe
 CC specific for a tumour rejection antigen precursor. Tumour rejection
 CC antigen precursors used in the present invention are the MAGE family,
 CC BAGE, GAGE, LAGE, NY-ESO-1 and PRAME (previously referred to as DAGB).
 CC Expression of the tumour rejection antigen precursor indicates possible
 CC multiple myeloma in the patient. The method can also be used for
 CC monitoring the disease progress and course of therapeutic regime. The
 CC present sequence is a peptide derived from a tumour rejection antigen
 CC precursor, which was used in the method of the present invention.

SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 22; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.00066;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLIMWITQCFL 11
 |||||
 1 SLIMWITQCFL 11

CC MAGE-A1 human leukocyte antigen (HLA)-B35 or HLA-B44 binding peptide,
CC or of a MAGE-A3 HLA-B35 binding peptide, identified by methods described
CC in the specification. MAGE genes encode tumor rejection antigens
CC (TRAs) presented to T lymphocytes by HLA-B35 and HLA-B44 molecules.
CC The MAGE antigenic peptide acts by binding to HLA molecules
CC on tumor cells and stimulating recognition of these cells and thus
CC signalling them to the immune system for destruction. The peptide when
CC presented by HLA molecule induces the activation and stimulation of
CC CD8+ cytotoxic T lymphocytes. The MAGE antigenic peptide is used to
CC treat and diagnose disorders characterised by expression of MAGE-A1
CC or A3. Disorders include cancers e.g. melanomas, oesophageal, lung,
CC head and neck, breast, colorectal, prostate, renal, bladder,
CC hepatocellular, papillary thyroid and gastric carcinomas, myelomas,
CC brain tumours, sarcomas, seminomas, and ovarian tumours. The present
CC sequence is human NY-ESO-1 tumour associated antigenic peptide
CC presented by major histocompatibility complex (MHC) HLA-A2. The
CC antigenic peptide is used in combination with peptides of the
CC invention for inducing an immune response.

SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 22; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00066;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLIMWITQCF 11
DB 1 SLIMWITQCF 11

RESULT 11

ID AAB85298 standard; peptide; 11 AA.

AC AAB85298;

DT 17-SEP-2001 (first entry)

DE HLA-A2 binding NY-ESO-1 peptide #1.

KW NY-ESO-1; human leukocyte antigen; HLA; lysis; cytolytic T cell; CTL;

KM HLA-A2; T-cell sorter; tumor; immune tetramer.

OS Homo sapiens.

PN MO200136453-A2.

PD 25-MAY-2001.

PF 08-NOV-2000; 2000MO-US42010.

PR 15-NOV-1999; 99US-0440621.

PR 25-FEB-2000; 2000US-0514036.

PR 29-SEP-2000; 2000US-0676005.

PA (LUDW-) LUDWIG INST CANCER RES.

PA (UYOX-) UNIV OXFORD.

PI Valmori D, Cerottini J, Romero P, Cerundolo V;

DR WPI; 2001-451454/48.

XX Novel isolated NY-ESO-1 nonapeptide useful for determining if a cell

XX presents human leukocyte antigen-A2 molecule on its surface, binds to

XX human leukocyte antigen molecules and provokes lysis by cytolytic T

XX cells -

XX Example 1; Page 4; 38pp; English.

XX The invention provides NY-ESO-1 peptide derivatives which bind to human
XX leukocyte antigen (HLA) molecules and provokes lysis by cytolytic T cells
XX (CTLs). The NY-ESO-1 nonapeptide is of formula SLIMWITQX, where X is an
XX amino acid having an uncharged polar side chain. The NY-ESO-1 peptide

CC derivatives are useful for determining if a cell presents an HLA-A2
CC molecule on its surface, by contacting a sample containing the cell with
CC the peptide or its derivative, and determining binding between them.
CC Where the binding is indicative of HLA-A2 on the surface of the cell.
CC The NY-ESO-1 peptides and analogues are useful therapeutically, for
CC administration to a patient who is HLA-A2 positive and expresses NY-ESO-1
CC in connection with the pathology, as well as diagnostically, i.e. to
CC determine if HLA-A2 positive cells are present, or if relevant CTLs are
CC present. They are also useful for determining the presence of CTLs in a
CC sample. The peptides are useful as T-cell sorters, when incorporated into
CC immune tetramers. The present sequence represents a NY-ESO-1 peptide that
CC can bind to HLA-A2 molecule.

SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 22; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00066;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLIMWITQCF 11
DB 1 SLIMWITQCF 11

RESULT 12

ID AAE02119 standard; peptide; 11 AA.

AC AAE02119;

DT 31-JUL-2001 (first entry)

DE NY-ESO-1 human leukocyte antigen-A2-binding peptide #1.

KW Human; cytostatic; immunogen; NY-ESO-1; human leukocyte antigen;

KM HLA; CD8; cytotoxic T lymphocyte; cancer; carcinoma; melanoma; myeloma;

KW brain tumour; sarcoma; vaccine; gene therapy.

OS Homo sapiens.

PN WO200129220-A2.

PD 26-APR-2001.

PF 19-OCT-2000; 2000MO-US28852.

PR 19-OCT-1999; 99US-0160374.

PR 01-FEB-2000; 2000US-0179570.

PA (LUDW-) LUDWIG INST CANCER RES.

PI Heidecker L, Van Den Eynde B, Boon-Falleur T, Brasseur F;

DR WPI; 2001-328498/34.

XX New antigenic peptides derived from MAGE-A12 polypeptides, useful for
XX diagnosis and treatment of cancer, such as bladder, lung, breast,
XX brain, prostate and renal carcinomas -

PS Disclosure; Page 21; 69pp; English.

XX The patent discloses antigenic peptides derived from MAGE-A12

XX protein and presented by human leukocyte antigens (HLAs). These

XX antigenic peptides when presented by an antigen presenting cell

XX having a HLA class I molecule, effectively induce the activation

XX and proliferation of CD8+ cytotoxic T lymphocytes (CTLs). MAGE-A12

XX is useful for treating a subject having a disorder characterised by

XX expression of MAGE-A12. The protein microarray comprising MAGE-A12

XX is useful for diagnosing a disorder, especially cancer, by determining

XX the binding of an antibody, T lymphocytes or a HLA molecule isolated

XX from the subject suspected of having the disorder characterised by the

XX expression of MAGE-A12. MAGE-A12 is useful for treating cancers,
XX including bladder carcinomas, melanomas, oesophageal, lung, head and


```
XX 22-FEB-2000; 2000US-0510635.
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Lethe B, Boon-Falleur T;
XX
XX WPI; 2001-550091/61.
XX
XX Genomic sequences of tumour associated antigen EY-ESO-1 (LAGS-2) useful
XX for diagnosing testicular tumours -
XX
XX Example 12; Page 24; 50pp; English.
XX
XX The present sequence represents a peptide which is derived from cancer
XX testis tumour antigen NY-ESO-1 (also called LAGS-2). The peptide
XX stimulates cytolytic T cell lines (CTLs). NY-ESO-1 is a molecule that
XX is processed to at least one human leukocyte antigen (HLA) binding
XX peptide, which binds to Class I and Class II major histocompatibility
XX complex (MHC). NY-ESO-1 is expressed in tumour mRNA and in testis,
XX but not normal colon, kidney, liver or brain tissue. The presence of
XX level of expression of NY-ESO-1 may be assayed for the diagnosis of
XX cancer, especially testis tumours.
XX
XX Sequence 11 AA;
XX
SQ
Query Match 100.0%; Score 61; DB 22; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00066;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLMMWITQCF 11
DB 1 SLMMWITQCF 11
RESULT 9
AAE07777
ID AAE07777 standard; peptide; 11 AA.
AC AAE07777;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human wildtype NY ESO-1 peptide, ESOP157-167.
XX
XX Human; major histocompatibility complex; MHC; vaccine; metastasis;
XX class II restricted T cell epitope; MHC-II epitope; cancer antigen;
XX NY ESO-1 protein; CD4+ T lymphocyte; human leucocyte antigen; HLA;
XX tumour-specific humoral-mediated immunity; cancer; cytostatic;
XX immunotherapy.
XX
XX Homo sapiens.
XX
XX WO200153833-A2.
XX
XX 02-AUG-2001.
XX
XX 26-JAN-2001; 2001WO-US02765.
XX
XX 28-JAN-2000; 2000US-0179004.
XX
XX 29-SEP-2000; 2000US-0237107.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Wang R, Rosenberg SA, Zeng G;
XX
XX WPI; 2001-496851/54.
XX
XX New NY-ESO cancer peptide or MHC class II restricted T cell epitopes,
XX useful as immunogen and vaccine for inhibiting cancer in a mammal or as
XX protection from metastasis -
XX
XX Example 14; Page 62; 134pp; English.
```

```
XX The invention relates to the identification and isolation of major
XX histocompatibility (MHC) class II restricted T cell epitope (MHC-II
XX epitope) derived from the cancer antigen, NY ESO-1. The MHC-II epitopes
XX from NY ESO-1 are recognised by CD4+ T lymphocytes in an human leucocyte
XX antigen (HLA) class II restricted manner, in particular HLA-DR or HLA-DP
XX restricted. The products of the gene are promising candidates for
XX immunotherapeutic strategies for the prevention, treatment and diagnosis
XX of patients with cancer. The cancer epitopes are useful as immunogen and
XX vaccine to inhibit or to prevent cancer in a mammal by eliciting CD4+ T
XX lymphocytes resulting in protection of the recipient from development of
XX cancer and protection from metastasis, or by inhibiting the growth of
XX cells expressing the NY-ESO-1 gene product. The cancer peptides are also
XX useful as diagnostic agent to detect the presence of cancer, to enhance
XX the generation of antibody and/or CD8+ T cell responses against any
XX given target antigen and/or happen and to induce tumour-specific
XX humoral-mediated immunity against cancer. The present sequence is human
XX wildtype NY ESO-1 peptide, ESOP157-167.
XX
XX Sequence 11 AA;
XX
SQ
Query Match 100.0%; Score 61; DB 22; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00066;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLMMWITQCF 11
DB 1 SLMMWITQCF 11
RESULT 10
AAE06849
ID AAE06849 standard; peptide; 11 AA.
AC AAE06849;
XX
XX 16-OCT-2001 (first entry)
XX
XX Human NY-ESO-1 antigenic peptide #1.
XX
XX MAGE antigenic peptide; Human leukocyte antigen; HLA-B35; HLA-B44;
XX tumour cell; immunostimulant; antigen presentation; cancer; melanoma;
XX CD8+ cytotoxic T lymphocyte; colorectal; prostate; gastric carcinoma;
XX myeloma; brain tumour; sarcoma; seminoma; ovarian tumour; cytostatic;
XX gene therapy; tumour rejection antigen; TRA; human; NY-ESO-1;
XX MHC; major histocompatibility complex.
XX
XX Homo sapiens.
XX
XX WO200153833-A1.
XX
XX 26-JUL-2001.
XX
XX 19-JAN-2001; 2001WO-US02808.
XX
XX 20-JAN-2000; 2000US-0177242.
XX
XX 25-OCT-2000; 2000US-0243212.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Luiten R, Boon-Falleur T, Van Der Bruggen P, Stroobant V;
XX Demotte N, Schultz E;
XX
XX WPI; 2001-488724/53.
XX
XX Functional variants and isolated mimetics of a MAGE-A1 HLA-B35 or
XX HLA-B44 binding peptide, or of a MAGE-A3 HLA-B35 binding peptide, used
XX in diagnosis and treatment of a disorder characterized by expression of
XX MAGE-A1 or -A3 -
XX
XX Disclosure; Page 28; 103pp; English.
XX
XX The invention relates to functional variants and isolated mimetics of a
```

```
XX OS Homo sapiens.
XX PN WO200000824-A1.
XX PD 06-JAN-2000.
XX PF 25-JUN-1999; 99WO-US14493.
XX PR 26-JUN-1998; 98US-0105639.
XX PA (LUDM-) LUDWIG INST CANCER RES.
XX PI Tureci O, Sahin U, Pfeundschn M, Rammensee G, Stevanovic S,
XX PI Chen Y, Gure A, Old LJ;
XX DR WPI; 2000-170933/15.
XX PT Determining the possible presence of breast, endometrial, colorectal,
XX PT lung, bladder or head-neck cancer -
XX PS Example 12; Page 21; 40pp; English.
XX CC A method has been developed for determining the possible presence of a
XX CC cancer, which is not melanoma or synovial sarcoma. The method comprises
XX CC assaying a sample taken from the subject to determine the expression of
XX CC an SSX gene, and determining the expression as a determination of the
XX CC possible presence of cancer. Expression of SSX1 gene indicates possible
XX CC presence of breast, endometrial, colorectal, lung, bladder or head-neck
XX CC cancer. These cancers are also detected by SSX2 and SSX4 gene
XX CC expression. SSX2 gene expression additionally indicates possible presence
XX CC of lymphoma, renal cell cancer, glioma and prostate cancer. Expression
XX CC of SSX4 gene also indicates possible presence of ovarian or stomach
XX CC cancer. SSX3 gene expression indicates the same cancers as SSX1, except
XX CC breast cancer. Determining expression of SSX gene can be used to monitor
XX CC progress of melanoma or synovial sarcoma, which is not cancer. The
XX CC SSX-derived peptide complex stimulates proliferation of cytolytic T
XX CC cells. This is useful for treating cancer, especially melanoma. AAY78464
XX CC to AAY78468 represent specifically claimed HLA binding peptides for use
XX CC in the method of the invention. AAY8452 to AAY8465 represent PCR
XX CC primers used in the isolation of SSX genes in the exemplification of the
XX CC present invention. AAY78469 to AAY78500, and AAY79684 to AAY79762
XX CC represent peptides derived from SSX proteins or NY-ESO-1, which are used
XX CC in the exemplification of the present invention.
XX SQ Sequence 11 AA;
XX QY Query Match 100.0%; Score 61; DB 21; Length 11;
XX QY Best Local Similarity 100.0%; Pred. No. 0.00066;
XX QY Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX DB 1 SLMMWITQCF 11
XX DB 1 SLMMWITQCF 11
XX RESULT 7
XX AAY52431
XX ID AAY52431 standard; peptide; 11 AA.
XX AC AAY52431;
XX XX
XX DT 15-FEB-2000 (first entry)
XX DE Human tumour antigen NY-ESO-1 peptide #4.
XX KM Cancer; tumour; antigen; MHC; major histocompatibility complex; Class I;
XX KM T-cell; cytotoxic; stimulation; proliferation; treatment;
XX KM diagnosis; prevention; melanoma; breast cancer; ovarian cancer;
XX KM prostate cancer; hepatoma; thyroid cancer; bladder cancer; lung cancer;
XX KM lymphoma.
XX OS Synthetic.
```

```
XX OS Homo sapiens.
XX PN WO9553938-A1.
XX PD 28-OCT-1999.
XX PF 24-MAR-1999; 99WO-US06875.
XX PR 17-APR-1998; 98US-0062422.
XX PR 02-OCT-1998; 98US-0165546.
XX PA (LUDM-) LUDWIG INST CANCER RES.
XX PI Stockert E, Jager E, Chen Y, Scanlan M, Alexander K, Old LJ;
XX PI Gure A, Ritter G;
XX DR WPI; 2000-038483/03.
XX PT Novel peptides which bind to MHC class I and MHC class II molecules,
XX PT useful for therapeutic and diagnostic purposes -
XX PS Claim 60; Page 18; 49pp; English.
XX CC Peptides #4-#7 (AAY52431-Y52434) are peptides derived from the human
XX CC tumour antigen, NY-ESO-1 (AAY52430) which contain the motif SLMMWIT
XX CC (AAY52441). These sequences can bind to MHC (major histocompatibility
XX CC Class I HLA-A2 molecules), thereby stimulating proliferation of
XX CC cytotoxic T-cells. cDNA encoding NY-ESO-1 was initially isolated
XX CC from an oesophagus squamous cell cancer cDNA library. Tissue
XX CC localisation studies revealed it to be expressed at high levels
XX CC in normal ovary and testis but not in normal colon, kidney, liver,
XX CC brain, oesophagus and skin. It was expressed in certain tumours and
XX CC tumour cell lines with some degree of frequency - these included
XX CC melanoma specimens and cell lines, and breast and bladder cancer.
XX CC specimens, with expression in other tumour types being sporadic.
XX CC These NY-ESO-1-derived peptides may be used in methods and
XX CC compositions used for the treatment, diagnosis and prevention of
XX CC cancers (such as melanoma, breast cancer, prostate cancer, lung
XX CC cancer, hepatoma, ovarian cancer, thyroid cancer, bladder cancer,
XX CC or lymphoma) and to stimulate the proliferation of T cells.
XX SQ Sequence 11 AA;
XX QY Query Match 100.0%; Score 61; DB 21; Length 11;
XX QY Best Local Similarity 100.0%; Pred. No. 0.00066;
XX QY Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX DB 1 SLMMWITQCF 11
XX DB 1 SLMMWITQCF 11
XX RESULT 8
XX AAG67165
XX ID AAG67165 standard; peptide; 11 AA.
XX AC AAG67165;
XX XX
XX DT 13-NOV-2001 (first entry)
XX DE Cancer testis tumour antigen NY-ESO-1 derived CTL-stimulating peptide.
XX KM Cancer testis tumour antigen; NY-ESO-1; LAGB-2; human leukocyte antigen;
XX KM HLA; HLA binding peptide; major histocompatibility complex; MHC;
XX KM tumour; cancer; testis tumour.
XX OS Homo sapiens.
XX PN WO200162917-A1.
XX PD 30-AUG-2001.
XX PF 22-JAN-2001; 2001WO-US02126.
```

XX	RESULT 4
XX	AAB22790
ID	AAB22790 standard; peptide; 11 AA.
XX	
AC	AAB22790;
DT	22-DEC-2000 (first entry)
DE	NY-ESO-1 peptide epitope, SEQ ID NO:1.
XX	
KM	NY-ESO-1; epitope; CTL response; cytotoxic T lymphocyte; vaccine;
KM	immunogenic; adjuvant coadministration; microbial infection;
KM	tuberculosis; HIV; hepatitis B virus; hepatitis C virus; cancer.
OS	Unidentified.
PN	MO20048630-AI.
PD	24-AUG-2000.
PF	17-FEB-2000; 2000WO-AU00110.
PR	17-FEB-1999; 99AU-0008735.
PR	27-JUL-1999; 99AU-0001861.
PA	(CSLUC-) CSL LTD.
EI	Cox JC, Drane DP;
DR	WPI; 2000-571930/53.
PT	Immunogenic complexes comprising negatively charged organic carrier
PT	adjuvants and positively charged antigens for use as vaccines against
PT	microbial infection and cancer -
PS	
XX	Example 4; Fig 5a; 11pp; English.
CC	The invention relates to a novel immunogenic complex comprising a
CC	charged organic carrier and a charged antigen which are
CC	electrostatically associated. The complex induces a cytotoxic T
CC	lymphocyte (CTL) response. The complex and/or vaccine can be used to
CC	treat a disease in a mammal, where the complex/vaccine elicits, induces
CC	or otherwise facilitates an immune response which inhibits, halts, delays
CC	or prevents the onset or progression of the disease condition. In
CC	particular, the disease is a condition resulting from a microbial
CC	infection or cancer. Microbial infections which may be treated using the
CC	immunogenic complex include human immunodeficiency virus (HIV), hepatitis
CC	B, hepatitis C, tuberculosis or a parasitic condition, and cancers which
CC	may be treated include melanoma, prostate cancer or breast cancer. The
CC	complexes and vaccines simultaneously co-deliver antigen and adjuvant to
CC	the same antigen presenting cell, which is often essential for induction
CC	of appropriate immune responses. Sequences AAB22790-B22791 represent
CC	peptide epitopes of the positively charged protein NY-ESO-1 used in an
CC	exemplification of the invention.
SQ	Sequence 11 AA:
Query Match	100.0%; Score 61; DB 21; Length 11;
Best Local Similarity	100.0%; Pred. No. 0.00066;
Matches 11; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 SLMMITOCFL 11
DB	1 SLMMITOCFL 11
RESULT 5	
AAB02630	
ID	AAB02630 standard; Peptide; 11 AA.
XX	
AC	AAB02630;

XX 18-AUG-2000 (first entry)
XX
XX Tumour associated peptide antigen from NY-ESO-1 #1.
DE
XX MAGE-A3; HLA class II; human leukocyte antigen; antibody; vaccine;
KW cancer; human; tumour; tumour associated gene product.
OS
XX Homo sapiens.
XX
XX WO200020581-A1.
XX
XX 13-APR-2000.
XX
XX 15-SEP-1999; 99WO-US21230.
XX
XX 05-OCT-1998; 98US-0166448.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX (UYVR-) UNIV VRIJE BRUSSEL.
XX
XX Chaux P, Strocobant V, Boon-Falleur T, Van Der Bruggen P;
PI Schulz BS, Van Snick J, Lethé B, Thielemans K, Corthals J;
PI Helrman C;
XX
XX WPI; 2000-317713/27.
XX
XX New MAGE-A3 class II binding peptides, useful to diagnose and treat
PT tumours, are fragments of MAGE-A3 which bind to and are presented to T
PT lymphocytes by human leukocyte antigen class II molecules -
XX
XX Disclosure; Page 33, 11pp; English.
XX
XX The present invention relates to MAGE-A3 (tumour associated gene
CC product) human leukocyte antigen (HLA) class II-binding peptides (see
CC AA802566-B02595, and AAB02633-B02637). These peptides are presented to T
CC cells in the context of HLA class II molecules. The peptides stimulate
CC the activity and proliferation of CD4+ T lymphocytes. The invention also
CC includes nucleotide sequences encoding MAGE-3A peptides (see ABA37928
CC and ABA37938-A37940). The peptides and nucleotide sequences can be used
CC to create antibodies against the MAGE-A3 peptides, the antibodies, the
CC peptides and nucleotide sequences can be used to create a vaccine. The
CC peptides are used to diagnose or treat a disorder characterized by
CC expression of MAGE-3, particularly cancer. The methods can also be used
CC in the diagnosis of disorders associated with MAGE-3 expression. Included
CC in the invention are other human tumour antigens (see AAB02596-B02637),
CC and PCR primers used in the course of the invention (see ABA37929-A37937
CC and ABA37941-A37942).
XX
XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 21; Length 11;
Best local Similarity 100.0%; Pred. No. 0.00066;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

CY 1 SLMMWITQCFL 11
1 SLMMWITQCFL 11
Db

RESULT 6
ID AAY78469
AC AAY78469; AA
XX
XX 10-MAY-2000 (first entry)
XX
XX NY-ESO-1 derived peptide #1.
DE
XX
XX Cancer; SSX family; SSX-1; SSX-2; SSX-3; SSX-4; SSX-5; NY-ESO-1;
KW HLA binding; human leukocyte antigen; cytolytic T cell; CTL;
KW cytotoxic; melanoma; synovial sarcoma.

[illegible]

CC	specific antibodies. Detection of of the peptides, e.g. in binding assays, particularly with antibodies, is used for diagnosis of such diseases.
CC	
XX	
XX	
SQ	Sequence 11 AA;
OY	Query Match 100.0%; Score 61; DB 20; Length 11; Best Local Similarity 100.0%; Pred. No. 0.00066; Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB	1 SLMTWITQCFL 11 1 SLMTWITQCFL 11
RESULT 3	
ID	AAB08702 standard; Peptide; 11 AA.
XX	
XX	AAB08702;
DT	02-JAN-2001 (first entry)
DE	Antigenic peptide from tumour rejection antigen NY-ESO-1.
KM	EphA3; HLA class II-binding peptide; human leukocyte antigen; antigen; CD4+ T lymphocyte; tumour associated gene; vaccine.
XX	
OS	Homo sapiens.
FN	WO300050589-A1.
PD	31-AUG-2000.
PE	18-FEB-2000; 2000WO-US04326.
BR	22-FEB-1999; 99US-0121170. 08-OCT-1999; 99US-0158566.
PA	(LUDW-) LUDWIG INST CANCER RES.
PI	Chiarl R, Coulie P, Boon-Falleur T;
DR	WPI; 2000-572089/53.
PT	Noval tyrosine kinase receptor, EphA3 human leukocyte antigen (HLA) class II binding peptide and nucleic acid encoding the receptor, useful for diagnosing and treating conditions characterized by expression of EphA3 gene -
PS	Disclosure; Page 35; 107pp; English.
XX	
AA08668-B08704	represent antigenic peptides characteristic of tumours. The peptides may be combined in vaccines with a human EphA3 HLA (human leukocyte antigen) class II-binding peptide. EphA3 antigens, when presented by an antigen presenting cell having a HLA class II molecule, effectively induce activation and proliferation of CD4+ T lymphocytes. EphA3 is a tumour associated gene. EphA3 HLA binding peptides are used for selectively enriching a population of T lymphocytes. The peptides are also used for diagnosing a disorder characterized by EphA3 or EphA3 HLA binding peptide expression. The peptides are also used to treat a disorder characterized by EphA3 expression. The EphA3 binding peptides are useful in producing vaccines and antibody.
SQ	Sequence 11 AA;
OY	Query Match 100.0%; Score 61; DB 21; Length 11; Best Local Similarity 100.0%; Pred. No. 0.00066; Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB	1 SLMTWITQCFL 11 1 SLMTWITQCFL 11

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 17:13:40 ; Search time 39.8276 Seconds

(without alignments)
43.839 Million cell updates/sec

Title: US-10-023-182-4
Perfect score: 61
Sequence: 1 SLMLWITQCFLL 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	61	100.0	11 19 AAM62585	Cancer associated
2	61	100.0	11 20 AA01761	Exemplary antigen
3	61	100.0	11 21 AAB08702	Antigenic peptide
4	61	100.0	11 21 AAB22790	NY-ESO-1 peptide e
5	61	100.0	11 21 AAB02630	Tumour associated
6	61	100.0	11 21 AA078469	NY-ESO-1 derived p
7	61	100.0	11 21 AAY52431	Human tumour anti
8	61	100.0	11 22 AAG67165	Cancer testis tunc
9	61	100.0	11 22 AAE07777	Human wildtype NY

10	61	100.0	11 22 AAE06849	Human NY-ESO-1 ant
11	61	100.0	11 22 AAB85298	HLA-A2 binding NY-
12	61	100.0	11 22 AAE02119	NY-ESO-1 human leu
13	61	100.0	11 22 AAU01536	Cytolytic T cell 1
14	61	100.0	11 22 AAB82016	HLA-A2 binding pep
15	61	100.0	11 22 AAB65947	Human NY-ESO-1 CTL
16	61	100.0	11 22 AAB31328	Exemplary antigen
17	61	100.0	11 24 ABE04812	Human NY-ESO-1 CTL
18	61	100.0	12 24 AAG66801	Tumour antigen NY-
19	61	100.0	14 22 AAE07758	Human NY-ESO-1 pro
20	61	100.0	14 22 AAE07764	Human HLA-DP restr
21	61	100.0	14 22 AAE07765	Human HLA-DP restr
22	61	100.0	14 22 AAE07766	Human HLA-DP restr
23	61	100.0	14 22 AAE07767	Human modified NY
24	61	100.0	20 22 AAE07768	Human modified NY
25	61	100.0	20 22 AAE07744	Human ESO p156-175
26	61	100.0	27 22 AAE07757	Human HLA-DP restr
27	61	100.0	28 24 AEP74371	Human NY-ESO-1 epi
28	61	100.0	30 23 AAU85112	Human NYNSOLA segm
29	61	100.0	30 23 AAU85128	Human LAGE1 segm
30	61	100.0	81 24 AEP72587	Melanoma poly-epit
31	61	100.0	81 24 AEP72588	Melanoma poly-epit
32	61	100.0	135 24 ABR48211	Human bladder canc
33	61	100.0	135 24 ABU56509	Lung cancer-associ
34	61	100.0	135 24 ABU56695	Lung cancer-associ
35	61	100.0	180 19 AAM65664	Human LAGE-1 clone
36	61	100.0	180 19 AAM65665	Human NY-ESO-1 pro
37	61	100.0	180 19 AAM62584	Cancer associated
38	61	100.0	180 20 AAY05965	Human cancer anti
39	61	100.0	180 21 AAB03154	Human oesophageal
40	61	100.0	180 21 AAY70860	Human LAGE-1 splic
41	61	100.0	180 21 AAY70862	Human tumour anti
42	61	100.0	180 21 AAY52430	Human tumour anti
43	61	100.0	180 22 AAG67164	Human tumour anti
44	61	100.0	180 22 AAE07714	Amino acid sequenc
45	61	100.0	180 22 AAU01535	Human NY-ESO-1 pro

ALIGNMENTS

```

RESULT 1
AAM62585
ID AAM62585 standard; peptide; 11 AA.
XX
AC AAM62585;
XX
DT 17-SEP-1998 (first entry)
XX
DE Cancer associated antigen peptide.
XX
KW Cancer associated antigen; NY-ESO-1; regression; progression; onset;
KW cancer; treatment; diagnosis.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN W09814464-A1.
XX
PD 09-APR-1998.
XX
PF 15-SEP-1997; 97WO-US16335.
XX
PR 03-OCT-1996; 96US-0725182.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Chen Y, Drifflout DW, Gure A, Jager E, Knuth A;
XX Old LJ, Scanlan M;
XX WPI; 1998-286417/25.
XX
PT New isolated cancer associated antigen - is used to develop products

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: February 10, 2004, 17:17:35 ; Search time 10.8621 Seconds
(without alignments)
79.683 Million cell updates/sec

Title: US-10-023-182-6

Perfect score: 46

Sequence: 1 QLSLWMIT 9

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	78.3	122	2	S54492
2	36	78.3	145	2	A64469
3	36	78.3	196	2	AE2876
4	36	78.3	198	2	G97652
5	35	76.1	142	2	B69515
6	35	76.1	228	2	S45677
7	35	76.1	311	2	T11362
8	35	76.1	349	2	B83434
9	35	76.1	368	2	T10557
10	34	73.9	674	2	A47222
11	34	73.9	676	2	B47222
12	34	73.9	806	2	T41930
13	34	72.9	1449	2	A12017
14	33	71.7	106	2	G81680
15	33	71.7	302	2	AE2520
16	33	71.7	305	2	S52972
17	33	71.7	385	2	G98024
18	33	71.7	787	2	T16901
19	33	71.7	1040	2	AD0721
20	33	71.7	3744	2	S46715
21	32	69.6	83	1	WSW1HS
22	32	69.6	121	1	K4HU
23	32	69.6	129	2	S40347
24	32	69.6	132	2	S46373
25	32	69.6	133	1	K4HU1
26	32	69.6	134	1	K4HU17
27	32	69.6	134	2	S21917
28	32	69.6	134	2	S49531
29	32	69.6	135	2	P86756

30	32	69.6	136	2	A49137	Ig kappa chain pre
31	32	69.6	138	2	A53261	Ig kappa chain pre
32	32	69.6	198	2	A71801	hypothetical prote
33	32	69.6	206	2	AE2222	hypothetical prote
34	32	69.6	240	2	S06084	Ig kappa chain pre
35	32	69.6	247	2	T52449	Bax inhibitor-1 li
36	32	69.6	278	2	G81139	hypothetical prote
37	32	69.6	308	2	S58995	NADH2 dehydrogenas
38	32	69.6	318	2	T11506	NADH2 dehydrogenas
39	32	69.6	390	2	T04096	glucose-6-phosphat
40	32	69.6	396	2	T04100	glucose-6-phosphat
41	32	69.6	407	2	S77146	hypothetical prote
42	32	69.6	456	1	I40516	spaf protein - Bac
43	32	69.6	463	2	AP0283	probable mannitol
44	32	69.6	534	2	AH2206	serine/threonine K
45	32	69.6	581	2	E69322	dolichol-P-glucose

ALIGNMENTS

RESULT 1
S54492
probable membrane protein YMR123w - Yeast (Saccharomyces cerevisiae)

N/Alternate names: hypothetical protein YM8564.05

C/Species: Saccharomyces cerevisiae

C/Date: 08-Jul-1995 #sequence_revision 17-Nov-1995 #text_change 19-Apr-2002

C/Accession: S54492

R/Lye, G.; Churcher, C.M.
submitted to the EMBL Data Library, May 1995

A/Reference number: S54014

A/Accession: S54492

A/Molecule type: DNA

A/Residues: 1-122 <LYE>

A/Cross-references: EMBL:Z49273; NID:G809577; PID:G809582; GSPDB:GN00013; MIPS:YMR123w

C/Genetics:

A/Gene: SGD:PKR1; MIPS:YMR123w

A/Cross-references: SGD:S0004730

A/Map position: 13R

C/Superfamily: Saccharomyces cerevisiae probable membrane protein YMR123w

C/Keywords: transmembrane protein

F/21-37/Domain: transmembrane #status predicted <TM>

F/51-67/Domain: transmembrane #status predicted <TM>

Query Match Best Local Similarity 78.3%; Score 36; DB 2; Length 122;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSLWMIT 9

DB 55 ISLLWMIT 62

RESULT 2

A64469
hypothetical protein MJ1354 - Methanococcus jannaschii

C/Species: Methanococcus jannaschii

C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

C/Accession: A64469

R/Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, J.; Reich, C.I.; Overbeek, R.; Kirsch, E.F.; Weissbrock, K.G.; Merrick, J.M.; Glodek, A.

rsen, J.D.; Sadow, P.W.; Haman, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A/Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, A.

A/Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A/Reference number: A64300; MUID:96337999; PMID:8688087

A/Accession: A64469

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-145 <BUL>

A/Cross-references: GB:U67575; GB:L77117; NID:G1591992; PDB:ABB9366.1; PID:G1591997;

C/Genetics:

A/Map position: FOR1304340-1304777

Query Match 78.3%; Score 36; DB 2; Length 145;
 Best Local Similarity 75.0%; Pred. No. 12;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSILMWT 9
 |||||:
 105 LSILMWT 112

RESULT 3

AE2876 conserved hypothetical protein Atu2439 [imported] - Agrobacterium tumefaciens (strain C5
 C/Species: Agrobacterium tumefaciens
 C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Dec-2002
 C/Accession: AE2876
 R/Wood, D.W.; Seubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutysavin, T.; Levy, R.; Li, M.; McClell
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C5.
 A/Reference number: AB2577; MUID:21608550; PMID:11743193
 A/Accession: AE2876
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-196 <KUR>
 A/Cross-references: GB:AE008688; PIDN:AL43427.1; PID:G17740928; GSPDB:GNO0186
 A/Experimental source: strain C58 (Dupont)
 C/Genetics:
 A/Gene: Atu2439
 A/Map position: circular chromosome
 C/Superfamily: uncharacterized conserved protein

Query Match 78.3%; Score 36; DB 2; Length 196;
 Best Local Similarity 85.7%; Pred. No. 16;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 SLIMWT 9
 |||||:
 23 SLIMWT 29

RESULT 4

G97652 hypothetical protein AGR_C_4424 [imported] - Agrobacterium tumefaciens (strain C58, cere
 C/Species: Agrobacterium tumefaciens
 C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Dec-2002
 C/Accession: G97652
 R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Onorlo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Marfeld, B.;
 Science 294, 2223-2238, 2001
 A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
 A/Reference number: A97359; MUID:21608551; PMID:11743194
 A/Accession: G97652
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-198 <KUR>
 A/Cross-references: GB:AE007869; PIDN:AAK8176.1; PID:G15157620; GSPDB:GNO0169
 C/Genetics:
 A/Gene: AGR_C_4424
 A/Map position: circular chromosome
 C/Superfamily: uncharacterized conserved protein

Query Match 78.3%; Score 36; DB 2; Length 198;
 Best Local Similarity 85.7%; Pred. No. 16;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 SLIMWT 9
 |||||:
 25 SLIMWT 31

Query Match 76.1%; Score 35; DB 2; Length 142;
 Best Local Similarity 77.8%; Pred. No. 18;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QSLIMWT 9
 |||||:
 47 QSLIMWT 55

RESULT 5

B69515 hypothetical protein AF2122 - Archaeoglobus fulgidus
 C/Species: Archaeoglobus fulgidus
 C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C/Accession: B69515
 R/Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 ; Fieschmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 ; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, D.
 Nature 390, 364-370, 1997
 A/Authors: Overbeek, R.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
 A/Reference number: A69250; MUID:98049343; PMID:9389475
 A/Accession: B69515
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-142 <KLE>
 A/Cross-references: GB:AE000958; GB:AE000782; NID:G2689281; PIDN:AB89149.1; PID:G264844

Query Match 76.1%; Score 35; DB 2; Length 142;
 Best Local Similarity 77.8%; Pred. No. 18;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QSLIMWT 9
 |||||:
 47 QSLIMWT 55

RESULT 6

S45677 proteinase inhibitor - signal crayfish
 C/Species: Pacifastacus lenisculus (signal crayfish)
 C/Date: 07-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 03-Feb-2003
 C/Accession: S45677
 R/Dohanson, M.W.; Keyser, P.; Soederhäll, K.
 Eur. J. Biochem. 223, 389-394, 1994
 A/Title: Purification and cDNA cloning of a four-domain Kazal proteinase inhibitor from
 A/Reference number: S45677; MUID:94333326; PMID:8055907
 A/Accession: S45677
 A/Molecule type: mRNA
 A/Residues: 1-228 <JOH>
 A/Cross-references: EMBL:X79512; NID:G498784; PIDN:CAA56043.1; PID:G498785
 A/Keywords: serine proteinase inhibitor
 F/21-66/Domain: Kazal proteinase inhibitor homology <KP1>
 F/21-119/Domain: Kazal proteinase inhibitor homology <KP2>
 F/123-170/Domain: Kazal proteinase inhibitor homology <KP3>
 F/174-221/Domain: Kazal proteinase inhibitor homology <KP4>

Query Match 76.1%; Score 35; DB 2; Length 228;
 Best Local Similarity 87.5%; Pred. No. 28;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LSILMWT 9
 |||||:
 3 LSILMWT 10

RESULT 7

T11362 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 1 - Daphnia pulex mitochondrion
 C/Species: mitochondrion Daphnia pulex
 C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
 C/Accession: T11362
 R/Cress, T.J.
 Gene 233, 89-99, 1999
 A/Title: The complete sequence of the mitochondrial genome of Daphnia pulex (Cladocera:
 A/Reference number: Z17264; MUID:99307147; PMID:10375625
 A/Accession: T11362
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-311 <CRB>

A:Cross-references: EMBL:AF117817; NID:g4927669; PID:g4927682; PIDN:AAD33242.1

C:Genetics:

A:Gene: NDI

A:Genome: mitochondrion

A:Genetic code: SGC4

C:Superfamily: NADH dehydrogenase (ubiquinone) chain 1

C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match

Best Local Similarity 76.1%; Score 35; DB 2; Length 311;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSLMWIT 9

Db 82 LSLMWLMS 89

RESULT 8

B83434

translocation protein in type III secretion PA1690 [imported] - Pseudomonas aeruginosa

C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: B83434

R:Stover, C.K.; Pham, X.Q.; Ewin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kae, A.; Lardig, K.; Lim,

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: A82950; PMID:20437377; PMID:10984043

A:Accession: B83434

A:Molecule type: DNA

A:Status: preliminary

A:Residues: 1-349 <STO>

A:Cross-references: GB:AE004596; GB:AE004091; NID:g9947658; PIDN:AA05079.1; GSPDB:GN001

C:Genetics:

A:Experimental source: strain PA01

A:Gene: pscY, PA1690

C:Superfamily: flagellar biosynthetic protein flhB; flhB carboxyl-terminal homology

Query Match

Best Local Similarity 76.1%; Score 35; DB 2; Length 349;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSLMWIT 9

Db 150 LSLMWLT 157

RESULT 9

T10557

hypothetical protein T12G13.140 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 15-Oct-1999

C:Accession: T10557

R:Bevan, M.; Lemard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro

submitted to the protein sequence database, June 1999

A:Reference number: Z16533

A:Accession: T10557

A:Molecule type: DNA

A:Residues: 1-368 <BEV>

A:Cross-references: EMBL:AL080252; GSPDB:GN00062; ATSP:T12G13.140

A:Experimental source: cultivar Columbia; BAC clone T12G13

C:Genetics:

A:Gene: ATSP:T12G13.140

A:Map position: 58/2; 79/2; 118/2; 201/3; 254/3; 305/2

Query Match

Best Local Similarity 76.1%; Score 35; DB 2; Length 368;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSLMWI 8

:|||||:

Db 211 ELSLMWT 218

RESULT 10

A47222

Kallmann syndrome protein homolog KAL - quail

C:Species: Phasianidae gen. sp. (quail)

C>Date: 24-Feb-1994 #sequence_revision 25-Apr-1997 #text_change 19-Feb-1999

C:Accession: A47222

R:Legouis, R.; Cohen-Salmon, M.; del Castillo, I.; Levlillers, J.; Caby, L.; Morrow, J

Genomics 17, 516-518, 1993

A:Title: Characterization of the chicken and quail homologues of the human gene respon

A:Reference number: A47222; PMID:94010957; PMID:8406507

A:Accession: A47222

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-674 <LEG>

A:Cross-references: GB:L13976; NID:g1196807; PID:g1196808

A:Note: sequence extracted from NCBI backbone (NCBIP:137995)

C:Superfamily: antileukoprotease repeat homology

P:124-170/Domain: antileukoprotease repeat homology <ALP>

Query Match

Best Local Similarity 73.9%; Score 34; DB 2; Length 674;

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSLMWIT 9

Db 10 LSLMWIT 17

RESULT 11

B47222

Kallmann syndrome protein homolog KAL - chicken

C:Species: Gallus gallus (chicken)

C>Date: 24-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 01-Dec-2000

C:Accession: B47222; A47394; S36170

R:Legouis, R.; Cohen-Salmon, M.; del Castillo, I.; Levlillers, J.; Caby, L.; Morrow, J

Genomics 17, 516-518, 1993

A:Title: Characterization of the chicken and quail homologues of the human gene respon

A:Reference number: A47222; PMID:94010957; PMID:8406507

A:Accession: B47222

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-676 <LEG>

A:Cross-references: GB:L12144; NID:g406510; PIDN:AA051435.1; PID:g406511

A:Note: sequence extracted from NCBI backbone (NCBIP:137996)

R:Legouis, R.; Lievre, C.A.; Leibovici, M.; Lapointe, F.; Petit, C.

Proc. Natl. Acad. Sci. U.S.A. 90, 2461-2465, 1993

A:Title: Expression of the KAL gene in multiple neuronal sites during chicken developm

A:Reference number: A47394; PMID:93211981; PMID:8460158

A:Accession: A47394

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 81-152, 'P', 154-237 <LEG>

A:Experimental source: embryo, olfactory bulb

A:Note: sequence extracted from NCBI backbone (NCBIP:128286; NCBIP:128287)

R:Rugart, E.I.; Lutz, B.; Kuratani, S.C.; Wawersik, S.; Borsani, G.; Ballabio, A.; Et

Nature Genet. 4, 19-26, 1993

A:Title: Expression pattern of the Kallmann syndrome gene in the olfactory system sugg

A:Reference number: S36170; PMID:93291868; PMID:8513320

A:Accession: S36170

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: WVR, 5-528, 'H', 530-576 <RUG>

C:Superfamily: antileukoprotease repeat homology

F:125-171/Domain: antileukoprotease repeat homology <ALP>

Query Match

Best Local Similarity 73.9%; Score 34; DB 2; Length 676;

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSLMWI 8

:|||||:

QY 2 LSLMWT 9
|:|:|:|:
Db 11 LALLMWT 18

RESULT 12

T41930
ribo-nucleotide reductase, large subunit - human herpesvirus 7 (strain J1)
C:Species: human herpesvirus 7
A:Variety: strain J1
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
R:Nicholas, J.
submitted to the EMBL Data Library, December 1995
A:Description: Determination and analysis of the complete nucleotide sequence of human H
A:Reference number: 222022
A:Accession: T41930
A:Status: preliminary; translated from GB/EMBL/DDDB
A:Molecule type: DNA
A:Residues: 1-806 <NIC>
A:Cross-references: EMBL:U43400; PIDN:AAC54690.1
A:Experimental source: strain J1
C:Genetics:
A>Note: U28

Query Match 73.9%; Score 34; DB 2; Length 806;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSLMWT 8
|:|:|:|:
Db 308 LSLMWT 314

RESULT 13

AI2017
hypothetical protein all1695 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A>Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AI2017
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi
Nakazaki, N.; Shimpō, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AI2017
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1449 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA578061.1; PID:G17135515; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all1695

Query Match 73.9%; Score 34; DB 2; Length 1449;
Best Local Similarity 87.5%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LSLMWT 9
|:|:|:|:
Db 1014 LSLMWT 1021

RESULT 14

G81680
hypothetical protein TC0636 [imported] - Chlamydia muridarum (strain N199)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MOPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C:Accession: G81680
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis MOPn and Chlamydia pneumoniae AR39.

A:Reference number: AB1500; MUID:20150255; PMID:10684935
A:Accession: G81680
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-106 <TET>
A:Cross-references: GB:AE002332; GB:AE002160; NID:97190670; PIDN:AAF39465.1; PID:G719067
A:Experimental source: strain N199 (MOPn)
C:Genetics:
A:Gene: TC0636

Query Match 71.7%; Score 33; DB 2; Length 106;
Best Local Similarity 75.0%; Pred. No. 31;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LSLMWT 9
|:|:|:|:
Db 80 LSLMWT 87

RESULT 15

AE2520
hypothetical protein alr7341 [imported] - Nostoc sp. (strain PCC 7120) plasmid PCC7120a1
C:Species: Nostoc sp. PCC 7120
A>Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AE2520
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi
Nakazaki, N.; Shimpō, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AE2520
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-302 <KUR>
A:Cross-references: GB:BA000020; PIDN:BA577099.1; PID:G17134540; GSPDB:GN00180
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr7341

Query Match 71.7%; Score 33; DB 2; Length 302;
Best Local Similarity 66.7%; Pred. No. 87;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QSLMWT 9
|:|:|:|:
Db 259 QSLMWT 267

Search completed: February 10, 2004, 17:29:47
Job time : 11.8621 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 17:13:40 ; Search time 6.82759 Seconds

(without alignments)
61.990 Million cell updates/sec

Title: US-10-023-182-6

Perfect score: 46
Sequence: 1 QLSLIMWIT 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	46	100.0	180	CTG1_HUMAN
2	36	78.3	122	YM10_YEAST
3	36	78.3	145	Y054_METTA
4	35	76.1	142	YL22_ARCFU
5	35	76.1	356	CKR8_MACMU
6	34	73.9	676	KALM_CHICK
7	34	72.9	806	RIR1_HSV7J
8	33	72.7	305	NYUW_AFLI
9	33	71.7	787	Y091_CAEBL
10	33	71.7	3744	YHP9_YEAST
11	32	69.6	83	VES_HPV16
12	32	69.6	121	KV40_HUMAN
13	32	69.6	133	KV4C_HUMAN
14	32	69.6	134	KV4C_HUMAN
15	32	69.6	167	YALU_RHISN
16	32	69.6	247	BIL_ARATH
17	32	69.6	308	NYUW_LUMTE
18	32	69.6	318	NYUW_PAPHA
19	32	69.6	463	DAUD_YERPE
20	32	69.6	512	ER24_SEBLY
21	32	69.6	794	CADC_HUMAN
22	32	69.6	1029	YK95_RHIME
23	31	67.4	120	VG19_BPMU
24	31	67.4	124	YBR3_YEAST
25	31	67.4	220	RADB_PYROO
26	31	67.4	301	NYUW_NYCN0
27	31	67.4	313	NYUW_LOCOM1
28	31	67.4	402	YBJT_ECOLI
29	31	67.4	437	PFTB_BOVIN
30	31	67.4	437	PFTB_HUMAN
31	31	67.4	437	PFTB_HUMAN
32	31	67.4	523	NYUW_PROW1
33	31	67.4	570	FKB9_MOUSE

34	31	67.4	578	1	AC22_STRCO	P46105 streptomyce
35	31	67.4	643	1	VE1_HPV27	P36723 human papill
36	31	67.4	643	1	VE1_HPV2A	P25481 human papill
37	31	67.4	643	1	VE1_HPV57	P22153 human papill
38	31	67.4	697	1	ATK6_LEPIN	P59219 leptospira
39	31	67.4	940	1	GBR2_RAT	O88871 rattus norv
40	31	67.4	1197	1	VGLM_RVFPV	P21401 rife valley
41	31	67.4	1206	1	VGLM_RVFPV	P03518 rife valley
42	31	67.4	1704	1	ABG3_HUMAN	O99758 homo sapien
43	31	67.4	1819	1	GCPE_HUMAN	O96677 homo sapien
44	30	65.2	186	1	YCF4_PORPU	P51220 porphyra pu
45	30	65.2	240	1	YCAN_PORPU	P51364 porphyra pu

ALIGNMENTS

RESULT 1
ID CTG1_HUMAN STANDARD; PRT; 180 AA.
AC P78358;
DT 15-DEC-1998 (Rel. 37, Last created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cancer/testis antigen 1 (Autoimmunogenic cancer/testis antigen NY-ESO-1).
DE CTG1 OR CTAG.
GN Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97203161; PubMed=9050879;
RA Chen Y.-T., Scailan M.J., Sahin U., Tureci O., Gure A.O., Tang S.,
RA Williamson B., Stockert B., Pfundschnub M., Old L.J.;
RT "A testicular antigen aberrantly expressed in human cancers detected
RT by autologous antibody screening."
RL Proc. Natl. Acad. Sci. U.S.A. 94:1914-1918(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Melanoma;
RX MEDLINE=98289662; PubMed=9626360;
RA Leche B., Lucas S., Michaux L., de Smet C., Godelaine D., Serrano A.,
RA de Plaen E., Boon T.;
RT "LAG-1, a new gene with tumor specificity."
RL Int. J. Cancer 76:903-908(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=98430682; PubMed=9759882;
RA Wang R.-F., Johnston S.L., Zeng G., Topalian S.L.,
RA Schwartzentruber D.J., Rosenberg S.A.;
RT "A breast and melanoma-shared tumor antigen: T cell responses to
RT antigenic peptides translated from different open reading frames."
RL J. Immunol. 161:3596-3606(1998).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN UTERINE MYOMETRIUM.
CC -1- SIMILARITY: BELONGS TO THE CTAG FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U87459; AAB49693.1; -
CC EMBL: AU003149; CAA05508.1; -
CC EMBL: AF038567; AAD05202.1; -
CC Genew: HGNC:2491; CTAG1.
CC MIM: 300156; -
CC Transmembrane, Antigen.

```

FT DOMAIN 5 82 GLY-RICH.
FT TRANSMEM 156 172 POTENTIAL.
SQ SEQUENCE 180 AA; 17992 MW; B122C5C2C8B81569 CRC64;

Query Match
Best Local Similarity 100.0%; Score 46; DB 1; Length 180;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLSLWMT 9
Db 155 QLSLWMT 163

RESULT 2
YML0_YEAST STANDARD; PRT; 122 AA.
ID YML0_YEAST
AC Q03880;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical 14.0 kDa protein in RPL15B-GCR3 intergenic region.
GN YML123W OR YML8564.05.
OC Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycaceae.
OX NCBI_TaxID=4932;
RN [1]
RP SQUENCE FROM N.A.
RC STRAIN=8288C / AB972;
RA PubMed=9269872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagsels K., Iye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrett B.G.;
RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XIII."
RL Nature 387:90-93 (1997).
CC -1- SIMILARITY: TO S.POMBE SPBP23A10.02.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z49273; CAA89272.1; -.
DR PIR; S54492; S54492.
DR SGD; S0004730; PKRI.
KM Hypothetical protein; Transmembrane.
FT TRANSMEM 21 41 POTENTIAL.
SQ SEQUENCE 122 AA; 13962 MW; 5CAA6929D55F2C40 CRC64;

Query Match
Best Local Similarity 75.0%; Score 36; DB 1; Length 122;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSLWMT 9
Db 55 LSLWMT 62

RESULT 3
YD54_METUA STANDARD; PRT; 145 AA.
ID YD54_METUA
AC Q58749;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein M1354.
GN M1354.

```

```

OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Meldrum J.F., Fuhrmann J.L., Nguyen D.,
RA Uitterlinden T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurt M.A., Kaine B.P., Bordovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073 (1996).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67575; AAB99366.1; -.
DR PIR; A64469; A64469.
DR TIGR; M1354; -.
KM Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 83 23 POTENTIAL.
FT TRANSMEM 103 103 POTENTIAL.
FT TRANSMEM 105 125 POTENTIAL.
SQ SEQUENCE 145 AA; 16713 MW; 15EFEE71C5262B37 CRC64;

Query Match
Best Local Similarity 75.0%; Score 36; DB 1; Length 145;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSLWMT 9
Db 105 LSLWMT 112

RESULT 4
YI22_ARCFU STANDARD; PRT; 142 AA.
ID YI22_ARCFU
AC O28158;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF2122.
GN AF2122.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwynn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.B., Kyrtides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uitterlinden T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,

```

RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon *Archaeoglobus fulgidus*.";
 RL Nature 390:364-370(1997).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF000958; AAB89149.1; -.
 DR PIR; B69515; B69515.
 DR TIGR; AF2122; -.
 KM Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 12 29 POTENTIAL.
 FT TRANSMEM 44 66 POTENTIAL.
 SQ SEQUENCE 142 AA; 15794 MW; B256B0C9A10048D CRC64;
 Query Match 76.1%; Score 35; DB 1; Length 142;
 Best Local Similarity 77.8%; Pred. No. 14;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 QSLMWIT 9
 Db 47 QSLMWIT 55
 RESULT 5
 CKR8_MACMU STANDARD; PRT; 356 AA.
 AC 097665;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE C-C chemokine receptor type 8 (C-C CKR-8) (CC-CKR-8) (CCR-8).
 GN CCR8.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=21354176; PubMed=1161684;
 RA Margalies B.J., Haer D.A., Clements J.E.;
 RT "Identification and comparison of eleven rhesus macaque chemokine
 RT receptors";
 RL AIDS Res. Hum. Retroviruses 17:981-986(2001).
 CC -1- FUNCTION: RECEPTOR FOR THE CHEMOKINES SCYLA/I-309, SCYAA/MIP-1-
 CC BETA AND SCYAL7/TARC. MAY REGULATE MONOCYTE CHEMOTAXIS AND THYMIC
 CC CELL LINE APOPTOSIS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF100205; AAC72403.1; -.
 DR InterPro; IPR004068; CC chemkin8.
 DR InterPro; IPR000276; GPCR_rhoopen.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PRO1530; CHEMOKIN8.
 DR PRINTS; PRO0237; GPCRHHODOPSN.

DR PROSITE; PS00237; G-PROTEIN RECP FL 1; 1.
 DR PROSITE; PS0262; G-PROTEIN RECP FL 2; 1.
 KM G-Protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 35
 FT TRANSMEM 1 63
 FT TRANSMEM 64 73
 FT DOMAIN 74 93
 FT TRANSMEM 94 107
 FT TRANSMEM 108 129
 FT DOMAIN 130 146
 FT TRANSMEM 147 172
 FT DOMAIN 173 203
 FT TRANSMEM 204 223
 FT DOMAIN 224 239
 FT TRANSMEM 240 264
 FT DOMAIN 265 281
 FT TRANSMEM 282 305
 FT DOMAIN 306 356
 FT DISULFD 106 184
 SQ SEQUENCE 356 AA; 41210 MW; 197628DEB4845B CRC64;
 Query Match 76.1%; Score 35; DB 1; Length 356;
 Best Local Similarity 75.0%; Pred. No. 33;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 2 LSLMWIT 9
 Db 154 LSLMWIT 161
 RESULT 6
 KALM_CHICK STANDARD; PRT; 676 AA.
 AC P33005;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Anosmin 1 precursor (Kallmann syndrome protein homolog).
 GN KAL.
 OS Gallus gallus (chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=94010957; PubMed=8406507;
 RA Legouis R., Cohen-Salmon M., del Castillo I., Leveilliers J.,
 RA Cayu L., Morrison J.-P., Petit C.;
 RT "Characterization of the chicken and quail homologues of the human
 RT gene responsible for the X-linked Kallmann syndrome";
 RL Genomics 17:516-518(1993).
 CC -1- FUNCTION: May be an adhesion-like molecule with anti-protease
 CC activity.
 CC -1- SUBCELLULAR LOCATION: Secreted. Localized at cell surface (By
 CC similarity).
 CC -1- SIMILARITY: Contains 4 fibronectin type III domains.
 CC -1- SIMILARITY: Contains 1 WAP-type domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L12144; AAA51435.1; -.
 DR PIR; B47222; B47222.
 DR HSPF; P19957; 2REL.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR002221; WAP.

```

DR Pfam; PF00041; fn3; 3.
DR Pfam; PF00095; wad; 1.
DR PRINTS; PR00003; 4DISULPHCORE.
DR SMART; SM00060; fn3; 3.
DR SMART; SM00217; WAP; 1.
DR PROSITE; PS00317; 4 DISULFIDE CORE; 1.
KW Cell adhesion; Glycoprotein; Serine protease inhibitor; Repeat;
Signal.
FT SIGNAL. 1 21 POTENTIAL.
FT CHAIN. 22 676 ANOSMIN 1.
FT DOMAIN. 22 115 "CRYSTEINE BOX".
FT DOMAIN. 125 171 WAP.
FT DOMAIN. 176 280 FIBRONECTIN TYPE-III 1.
FT DOMAIN. 281 397 FIBRONECTIN TYPE-III 2.
FT DOMAIN. 398 536 FIBRONECTIN TYPE-III 3.
FT DOMAIN. 537 657 FIBRONECTIN TYPE-III 4.
FT CARBOHYD. 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD. 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD. 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD. 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD. 549 549 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD. 560 560 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 676 AA; 76375 MW; 3FAC7ED82EA7E352 CRC64;

Query Match 73.9%; Score 34; DB 1; Length 676;
Best Local Similarity 62.5%; Pred. No. 90;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSLMWIT 9
DB 11 LALLLWMT 18

RESULT 7
RIRL HSV7J STANDARD; PRT; 806 AA.
AC P50641;
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribonucleoside-diphosphate reductase large chain (EC 1.17.4.1)
DE (Ribonucleotide reductase).
GN U28.
OS Human herpesvirus (type 7 / strain JI) (HHV7).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=57278;
RN [1]
RP SEQUENCE FROM N.A.
RA Nicholas J.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
thioedoxin + H(2)O = ribonucleoside diphosphate + reduced
thioedoxin.
CC -1- PATHWAY: DNA replication pathway; first step.
CC -1- SUBUNIT: Heterodimer of a large and a small chain.
CC -1- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
LARGE CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U43400; AAC54690.1; -.
DR PIR; T41930; T41930.
DR InterPro; IPR000788; Ribonucleo_red.
DR Pfam; PF00317; ribonuc_red_1g; 1.
DR Pfam; PF02867; ribonuc_red_1gc; 1

```

```

DR PROSITE; PS00089; RIBORED_LARGE; 1.
DR Oxidoreductase; DNA replication.
SQ SEQUENCE 806 AA; 93069 MW; 4BA516084FCCDC0 CRC64;

Query Match 73.9%; Score 34; DB 1; Length 806;
Best Local Similarity 85.7%; Pred. No. 1,1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSLMWIT 8
DB 308 LSLMWIT 314

RESULT 8
NUM APILI STANDARD; PRT; 305 AA.
ID NUM APILI
AC P34847;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 30-MAY-2000 (Rel. 32, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 1 (EC 1.6.5.3).
GN ND1.
OS Apis mellifera ligustica (Common honeybee).
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preygora;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae; Apis.
OX NCBI_TaxID=7469;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-Thorax;
RX MEDLINE=93114603; PubMed=8417993;
RA Crozier R.H., Crozier Y.C.;
RL "The mitochondrial genome of the honeybee Apis mellifera: complete
sequence and genome organization."
RT Genetics 133:97-117(1993).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L06178; AAB96810.1; -.
DR PIR; S52972; S52972.
DR InterPro; IPR001694; Resp_NADH_dhl.
DR Pfam; PF00146; NADHdh; 1.
DR PROSITE; PS00667; COMPLEX1_ND1_1; FALSE NEG.
DR PROSITE; PS00668; COMPLEX1_ND1_2; FALSE NEG.
DR Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
KW SEQUENCE 305 AA; 36838 MW; 01C4F46DCF7A6B6 CRC64;

Query Match 71.7%; Score 33; DB 1; Length 305;
Best Local Similarity 85.7%; Pred. No. 65;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSLMWIT 8
DB 80 LSLWMTI 86

RESULT 9
Y091 CAEEL
ID Y091 CAEEL STANDARD; PRT; 787 AA.
AC P41842;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein T30B3.2.1 in chromosome III

```

```

GN T20B12.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OK NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Brictol N2;
RA Waterston R.;
RL Submitted (JUN-1994) to the EMBL/GenBank/DBSJ databases.
CC -1- SIMILARITY: TO YEAST YNL313C.
CC -1- SIMILARITY: Contains 6 TPR repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, U10401; AAA19054.3; -.
CC PIR, T16901; T16901.
CC WormPep: T20B12.1; CE01407.
CC InterPro: IPR001440; TPR.
CC Pfam, PF00515; TPR; 4.
CC SMART, SM00028; TPR; 4.
CC Hypothetical protein; Repeat; TPR repeat.
CC REPEAT 405 432 TPR 1.
CC REPEAT 433 465 TPR 2.
CC REPEAT 467 500 TPR 3.
CC REPEAT 501 534 TPR 4.
CC REPEAT 536 568 TPR 5.
CC REPEAT 569 602 TPR 6.
CC SEQUENCE 787 AA; 89957 MW; 8C954156B3E85D52 CRC64;

Query Match 71.7%; Score 33; DB 1; Length 787;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QLSLMT 8
DB 277 QLSLMT 284

RESULT 10
YHP9_YEAST STANDARD; PRT; 3744 AA.
ID YHP9_YEAST
AC P38811;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 433.2 kDa protein in HXT5-NKX1 intergenic region.
GN YHR099M.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OK NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C/AB972;
MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favell A., Fulton L., Gatling S., Giesel C., Kirsten J.,
RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
RA Latreille P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,
RA Nhan M., Rifkin L., Riles L., St Peter H., Trevisan S., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VII."
RL Science 265:2077-2082(1994).
-1- SIMILARITY: STRONG, TO S.POWBE SPAC1F5.11C.

```

```

CC -1- SIMILARITY: BELONGS TO THE PI3/P14-KINASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, U00060; AAB68923.1; -.
CC PIR, S46715; S46715.
CC SGD, S0001141; YHR099M.
CC GO, GO:0000123; C:histone acetyltransferase complex; IPI.
CC GO, GO:0004402; P:histone acetyltransferase activity; IDA.
CC GO, GO:0016573; P:histone acetylation; IDA.
CC GO, GO:0006357; P:regulation of transcription from Pol II pro. .; IPI.
CC InterPro: IPR003157; FAT.
CC InterPro: IPR003152; FATC.
CC InterPro: IPR000403; PI3_P14_kinase.
CC InterPro: IPR001440; TPR.
CC Pfam, PF02259; FAT; 1.
CC Pfam, PF02260; FATC; 1.
CC Pfam, PF00454; PI3_P14_kinase; 1.
CC SMART, SM00146; PI3Kc_1.
CC Hypothetical protein; Transferase; Kinase.
CC DOMAIN 3414 3744 PI3K/P14K.
CC SEQUENCE 3744 AA; 433171 MW; AB3588676F5D5777 CRC64;

Query Match 71.7%; Score 33; DB 1; Length 3744;
Best Local Similarity 55.6%; Pred. No. 6.6e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 QLSLMT 9
DB 2822 QLSLMT 2830

RESULT 11
VES_HPV16 STANDARD; PRT; 83 AA.
ID VES_HPV16
AC P06927;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Probable E5 protein.
GN E5.
OS Human papillomavirus type 16.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OK NCBI_TaxID=10581;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=88160059; PubMed=2831662;
RA Budd V., McCance D.J., Schlegel R.;
RT "DNA sequence of the HPV-16 E5 ORF and the structural conservation of
RT its encoded protein."
RL Virology 163:243-246(1988).
RN [2]
RP SEQUENCE OF 20-83 FROM N.A.
MEDLINE=85246220; PubMed=2990099;
RA Seedorf K., Kramer G., Durr M., Sunat S., Roweckamp W.G.;
RT "Human papillomavirus type 16 DNA sequence."
RL Virology 145:181-185(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

DR EMBL; K02718; AAA46938.1; ALT_SEQ.
 DR PIR; A30016; MSMLS.
 DR InterPro; IPR004270; Papilloma_E5.
 DR Pfam; PF03025; Papilloma_E5; 1.
 KM Early protein.
 SQ SEQUENCE 83 AA; 9401 MW; 442C0ABFDD77C64 CRC64;

Query Match 69.6%; Score 32; DB 1; Length 83;
 Best Local Similarity 75.0%; Pred. No. 29;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ISLWMIT 9
 DB 45 LVLLWMT 52

RESULT 12
 KVA0 HUMAN STANDARD; PRT; 121 AA.

AC P06312;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ig kappa chain V-IV region precursor (Fragment).
 GN IGKV4-1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86041853; PubMed=2997712;
 RA Klobbeck H.G., Bornkamm G.W., Combratio G., Mociak R., Pohlenz H.D.,
 RA Zachau H.G.;
 RT "Subgroup IV of human immunoglobulin K light chains is encoded by a
 RT single germline gene";
 RL Nucleic Acids Res. 13:6515-6529(1985).
 CC -1- MISCELLANEOUS: THERE IS ONLY ONE IG KAPPA V-IV GENE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z00023; CAA77318.1; -.
 DR PIR; A01902; KAHU.
 DR HSSP; P80362; 1WTL.
 DR Genew; HGNC:5834; IGKV4-1.
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0003823; F:antigen binding activity; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_V.
 DR Pfam; PF00047; Ig_1.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS50835; IG-Like; 1.
 DR Immunoglobulin V region; Signal.
 FT CHAIN 1
 FT SIGNAL 20
 FT DOMAIN 21 >121 IG KAPPA CHAIN V-IV REGION.
 FT DOMAIN 21 43 FRAMEWORK-1.
 FT DOMAIN 44 60 COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 61 75 FRAMEWORK-2.
 FT DOMAIN 76 82 COMPLEMENTARITY-DETERMINING-2.
 FT DOMAIN 83 114 FRAMEWORK-3.
 FT DOMAIN 115 121 COMPLEMENTARITY-DETERMINING-3.
 FT DISULFID 43 114 BY SIMILARITY.
 FT NON_TER 121 121
 SQ SEQUENCE 121 AA; 13380 MW; 9586AD4189D33974 CRC64;

Query Match 69.6%; Score 32; DB 1; Length 121;
 Best Local Similarity 62.5%; Pred. No. 41;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 ISLWMIT 9
 DB 9 ISLWMTS 16

RESULT 13
 KVA8 HUMAN STANDARD; PRT; 133 AA.

AC P06313;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig kappa chain V-IV region JI precursor.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86041853; PubMed=2997712;
 RA Klobbeck H.G., Bornkamm G.W., Combratio G., Mociak R., Pohlenz H.D.,
 RA Zachau H.G.;
 RT "Subgroup IV of human immunoglobulin K light chains is encoded by a
 RT single germline gene";
 RL Nucleic Acids Res. 13:6515-6529(1985).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z00022; CAA77317.1; -.
 DR PIR; A01904; KAHUJ.
 DR HSSP; P80362; 1WTL.
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0003823; F:antigen binding activity; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_V.
 DR Pfam; PF00047; Ig_1.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS50835; IG-Like; 1.
 DR Immunoglobulin V region; Signal.
 FT CHAIN 1
 FT SIGNAL 20
 FT DOMAIN 21 133 IG KAPPA CHAIN V-IV REGION JI.
 FT DOMAIN 21 43 FRAMEWORK-1.
 FT DOMAIN 44 60 COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 61 75 FRAMEWORK-2.
 FT DOMAIN 76 82 FRAMEWORK-3.
 FT DOMAIN 83 114 COMPLEMENTARITY-DETERMINING-2.
 FT DOMAIN 115 122 COMPLEMENTARITY-DETERMINING-3.
 FT DISULFID 43 114 BY SIMILARITY.
 FT NON_TER 121 133
 SQ SEQUENCE 133 AA; 14632 MW; 5FB3953066744AF4 CRC64;

Query Match 69.6%; Score 32; DB 1; Length 133;
 Best Local Similarity 62.5%; Pred. No. 45;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 ISLWMIT 9
 DB 9 ISLWMTS 16

```

RESULT 14
KV4C_HUMAN STANDARD; PRT; 134 AA.
ID P06314;
AC 01-JAN-1988 (Rel. 06, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE 19 kappa chain V-IV region B17 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86011854; PubMed=2997713;
RA Marsh P., Mills F., Gould H.;
RT "Detection of a unique human V kappa IV germline gene by a cloned
RL cDNA probe."
RL Nucleic Acids Res. 13:6531-6544(1985).
RN [2]
RP REVISION TO 76.
RA Marsh P.;
RL Submitted (OCT-1986) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02990; CAA26733.1; -
DR HSSP; P80362; 1MTL.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; P:antigen binding activity; NAS.
DR GO; GO:0005955; P:immune response; NAS.
DR InterPro; IPR007110; IG_Like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; IG_1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PSS0835; IG_Like; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 134 IG KAPPA CHAIN V-IV REGION B17.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 44 60 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 61 75 FRAMEWORK-2.
FT DOMAIN 76 82 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 83 114 FRAMEWORK-3.
FT DOMAIN 115 121 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 122 133 FRAMEWORK-4.
FT DISULFID 43 114 BY SIMILARITY.
FT NON TER 134 134
SQ SEQUENCE 134 AA; 14966 MW; 6413A23FDD0738832 CRC64;

Query Match 69.6%; Score 32; DB 1; Length 134;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 ISLIMMIT 9
DB 9 ISLIMMIT 16

RESULT 15
Y4LJ_RHISN STANDARD; PRT; 167 AA.
ID Y4LJ_RHISN
AC P55550;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)

```

```

DE Hypothetical 18.1 kDa protein Y4LJ.
GN Y4LJ.
OS Rhizobium sp. (strain NGR234).
OG Plasmid sym PNGR234.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.T., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes."
RL Nature 387:394-401(1997).
CC -1- SIMILARITY: NONE OBVIOUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000083; AAB91762.1; -
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 167 AA; 18146 MW; 22F5D0340142E602 CRC64;

Query Match 69.6%; Score 32; DB 1; Length 167;
Best Local Similarity 66.7%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QLSIMMIT 9
DB 111 QLSIMMIT 119

```

Search completed: February 10, 2004, 17:18:06
 Job time : 7.82759 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 17:14:10 ; Search time 26.069 Seconds
(without alignments)
89.090 Million cell updates/sec

Title: US-10-023-182-6

Perfect score: 46

Sequence: 1 QLSLMWIT 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	180	4 Q9Y479	Q9Y479 homo sapien
2	38	82.6	401	11 Q8C0T0	Q8C0T0 mus musculi
3	36	78.3	198	16 Q80CP4	Q80CP4 agrobacteri
4	36	78.3	384	11 Q921M9	Q921M9 mus musculi
5	35	76.1	122	8 Q8HH81	Q8HH81 diglymma c1
6	35	76.1	130	8 Q8HH80	Q8HH80 diglymma c1
7	35	76.1	132	8 Q8HH79	Q8HH79 diglymma c1
8	35	76.1	228	5 Q26057	Q26057 pacifastacu
9	35	76.1	251	6 Q8KE76	Q8KE76 chlorobium
10	35	76.1	311	8 Q9TC86	Q9TC86 daphnia pul
11	35	76.1	349	16 Q91337	Q91337 pseudomonas
12	35	76.1	368	10 Q9SUFO	Q9SUFO arabidopsis
13	35	76.1	717	12 Q41942	Q41942 murid herpe
14	35	76.1	739	10 Q817F7	Q817F7 arabidopsis
15	34	73.9	64	16 Q8FKZ7	Q8FKZ7 escherichia
16	34	73.9	134	16 Q8F6X7	Q8F6X7 leptospira

17	34	73.9	156	16 Q9CLUB	Q9CLUB pasteurrella
18	34	73.9	202	5 Q95YB6	Q95YB6 caenorhabdi
19	34	73.9	266	2 Q9AEF8	Q9AEF8 fluoribacte
20	34	73.9	322	8 Q94T54	Q94T54 zu cristatu
21	34	73.9	322	8 Q94T67	Q94T67 trachipteru
22	34	73.9	378	16 Q9CM65	Q9CM65 pasteurrella
23	34	73.9	613	5 Q8IDF3	Q8IDF3 plasmodium
24	34	73.9	620	17 Q66XC8	Q66XC8 sulfobolus
25	34	73.9	674	13 Q90369	Q90369 coturnix co
26	34	73.9	675	13 Q9PSH7	Q9PSH7 gallus gall
27	34	73.9	750	16 Q8KDT8	Q8KDT8 chlorobium
28	34	73.9	1449	16 Q8YWC0	Q8YWC0 anabaena sp
29	34	73.9	89	5 Q9M4I6	Q9M4I6 drosophila
30	33	71.7	106	16 Q9PK37	Q9PK37 chlamydia m
31	33	71.7	108	17 Q96X80	Q96X80 sulfobolus
32	33	71.7	109	8 Q79150	Q79150 lymantria m
33	33	71.7	302	16 Q8YKF7	Q8YKF7 anabaena sp
34	33	71.7	338	8 Q8SKU0	Q8SKU0 aphelenchoi
35	33	71.7	385	16 Q8DPD8	Q8DPD8 streptococc
36	33	71.7	458	16 Q8FEC2	Q8FEC2 escherichia
37	33	71.7	541	10 Q8RUV6	Q8RUV6 oryza sativ
38	33	71.7	558	13 Q9PVE5	Q9PVE5 heterodontu
39	33	71.7	713	10 Q9FGS9	Q9FGS9 arabidopsis
40	33	71.7	728	10 Q8VYI3	Q8VYI3 arabidopsis
41	33	71.7	739	13 P79908	P79908 salmo salar
42	33	71.7	739	13 Q9PT31	Q9PT31 oncomychnu
43	33	71.7	749	10 Q9LE23	Q9LE23 oryza sativ
44	33	71.7	778	13 Q8UVV2	Q8UVV2 ginglymosco
45	33	71.7	896	10 Q8GSM2	Q8GSM2 hordeum vul

ALIGNMENTS

RESULT 1

Q9Y479 PRELIMINARY; PRT; 180 AA.

AC Q9Y479; 01-NOV-1999 (TREMURel. 12, Created)
DT 01-NOV-1999 (TREMURel. 12, last sequence update)
DT 01-OCT-2002 (TREMURel. 22, last annotation update)
DE LAGE-1S protein (Cancer/testis antigen 2).
GN LAGE1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=99325550; PubMed=10399963;
RA Aarnoudse C.A., Van den Boel P.B., Heemskerk B., Schrier P.I.,
RT "Interleukin-2-induced, melanoma-specific T cells recognize CAMEL, an
RT unexpected translation product of LAGE-1.";
RL Int. J. Cancer 82:442-448(1999).
[2]
RN SEQUENCE FROM N.A.
RA Aradhya S., Bardaro T., Galgoczy P., Yamagata T., Esposito T.,
RA Patlan H., Ciccodicola A., Kenwright S., Platzer M., D'Urso M.,
RA Nelson D.L.;
RT "Multiple pathogenic and benign genomic rearrangements occur at a 35-
RT kb duplication involving the NEMO and the LAGE2 genes.";
RL Hum. Mol. Genet. 0:0-0(2001).
DR EMBL; AJ012834; CA01094.1; -;
DR EMBL; AF272315; AA22015.1; -;
SQ SEQUENCE 180 AA; 18236 MW; 9077FA953543A25 CRC64;

Query Match 100.0%; Score 46; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q9Y479 1 QLSLMWIT 9
|||||||

Db 155 QSLIMWIT 163

RESULT 2

Q8COTO PRELIMINARY; PRT; 401 AA.
 ID Q8COTO;
 AC Q8COTO;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical type-1 copper.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 DR EMBL; AK029920; BAC26677.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 401 AA; 45562 MW; F34F237653D6FCF9 CRC64;

Query Match 82.6%; Score 38; DB 11; Length 401;
 Best Local Similarity 75.0%; Pred. No. 34;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSLIMWIT 9
 Db 315 ISLIMWIT 322

RESULT 3
 Q8UCP4 PRELIMINARY; PRT; 198 AA.
 ID Q8UCP4;
 AC Q8UCP4;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical protein Acu2439.
 GN ATU2439 OR AGR_C4424.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium.
 NX NCBI_TaxId=176239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Seubert J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D. Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kuyavayn T., Levy R., Li M.-J., McClelland E., Palmeri A.,
 RA Raymond C., Rouse Z., Saenphimachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tinney S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.,
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 RT C58."
 RL Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Hallin C., Mullin L.,
 RA Hummel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
 RA Woliam C., Allinger M., Doughty D., Scott C., Lapas C., Markelz B.,

RA Planagan C., Crowell C., Guron J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.; the plant pathogen and biotechnology agent
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58."
 RL Science 294:2323-2328(2001).
 DR EMBL; AE009191; AAL43427.1; ALT_INIT.
 DR EMBL; AE008157; AAK88176.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 198 AA; 21154 MW; FA40C6E548F0562F CRC64;

Query Match 78.3%; Score 36; DB 16; Length 198;
 Best Local Similarity 85.7%; Pred. No. 43;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 SLIMWIT 9
 Db 25 SLIMWIT 31

RESULT 4
 Q92IM9 PRELIMINARY; PRT; 384 AA.
 ID Q92IM9;
 AC Q92IM9;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Unknown (Protein for IMAGE:3709937) (Fragment).
 GN 903002E12R1K OR A4409659.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC011340; AAH11340.1; -.
 DR MGD; MGI:1925928; 903002E12R1K.
 FT NON_TER
 SQ SEQUENCE 384 AA; 42406 MW; 4B5F17D30FE16080 CRC64;

Query Match 78.3%; Score 36; DB 11; Length 384;
 Best Local Similarity 66.7%; Pred. No. 77;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QSLIMWIT 9
 Db 201 QSLIMWIT 209

RESULT 5
 Q8HH81 PRELIMINARY; PRT; 122 AA.
 ID Q8HH81;
 AC Q8HH81;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE NADH dehydrogenase subunit 1 (Fragment).
 GN NDI.
 OS Diglymma clyvioides.
 OS Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Coleoptera; Adephaga; Carabidae;
 OC Carabidae; Conjuridae; Diglymma.
 NX NCBI_TaxId=185701;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1;
 RA Rawson S.M., Emberson R., Armstrong K., Paterson A.;
 RT "Revision of the New Zealand endemic ground beetle genus Oregus
 RT Putney's 1868 (Carabidae: Broscini)."
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF466875; AAC14621.1; -.

KW Mitochondrion.
 FT NON TER 122 122
 SQ SEQUENCE 122 AA; 14223 MW; 43BED39602FD94DF CRC64;

Query Match
 Best Local Similarity 76.1%; Score 35; DB 8; Length 122;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LSLIMWI 8
 |||||
 DB 88 LSLIMWI 94

RESULT 6

OQHH80 PRELIMINARY; PRT; 130 AA.

AC OQHH80;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 GN NADH dehydrogenase subunit 1 (Fragment).
 OS Diglymma clivinoidea.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Coleoptera; Adephaga; Carabidae;
 CC Carabidae; Conjuridae; Diglymma.
 CX NCBI_TaxID=185701;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2;
 RA Pawson S.M., Emberson R., Armstrong K., Paterson A.;
 RT "Revision of the New Zealand endemic ground beetle genus Oregus
 Purveys 1968 (Carabidae: Broscini).";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF46876; AAO14622.1; --
 KW Mitochondrion.
 FT NON TER 130 130
 SQ SEQUENCE 130 AA; 15055 MW; 36FF2B6EDE2363E4 CRC64;

Query Match
 Best Local Similarity 76.1%; Score 35; DB 8; Length 130;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LSLIMWI 8
 |||||
 DB 88 LSLIMWI 94

RESULT 7
 OQHH79 PRELIMINARY; PRT; 132 AA.

AC OQHH79;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 GN NADH dehydrogenase subunit 1 (Fragment).
 OS Diglymma clivinoidea.
 OG Mitochondrion.
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Coleoptera; Adephaga; Carabidae;
 CC Carabidae; Conjuridae; Diglymma.
 CX NCBI_TaxID=185701;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=3;
 RA Pawson S.M., Emberson R., Armstrong K., Paterson A.;
 RT "Revision of the New Zealand endemic ground beetle genus Oregus
 Purveys 1968 (Carabidae: Broscini).";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF46877; AAO14623.1; --
 KW Mitochondrion.

FT NON TER 132 132
 SQ SEQUENCE 132 AA; 15257 MW; 2DE0D6FF2B6EDE23 CRC64;

Query Match
 Best Local Similarity 76.1%; Score 35; DB 8; Length 132;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LSLIMWI 8
 |||||
 DB 88 LSLIMWI 94

RESULT 8

OQ6057 PRELIMINARY; PRT; 228 AA.

AC OQ6057;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 GN PAPI I protein.
 OS Pacificastacus leniusculus (Signal crayfish).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 CC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
 CC Astacoidae; Astacidae; Pacificastacus.
 CX NCBI_TaxID=6720;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9433326; PubMed=8055907;
 RA Johansson M.W., Keyser P., Soederhaell K.;
 RT "Purification and cDNA cloning of a four-domain Kazal proteinase
 RT inhibitor from crayfish blood cells.";
 RL Eur. J. Biochem. 223:389-394 (1994).
 DR EMBL; X79512; CAA56043.1; --
 DR HSSP; P01004; 3SGE.
 DR InterPro; IPR002350; kazal.
 DR InterPro; IPR001239; kazal_inhib.
 DR Pfam; PF00050; kazal; 4.
 DR PRINTS; PRO0290; KAZALINHTR.
 DR SMART; SM00280; KAZAL; 4.
 KW Proteinase inhibitor, Serine protease inhibitor.
 FT CHAIN 20
 SQ SEQUENCE 228 AA; 24887 MW; DDCP3BFE8741747C CRC64;

Query Match
 Best Local Similarity 76.1%; Score 35; DB 5; Length 228;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 LSLIMWI 9
 |||||
 DB 3 LSLIMWI 10

RESULT 9
 OQKE76 PRELIMINARY; PRT; 251 AA.

AC OQKE76;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 GN Hypothetical protein CT0814.
 OS Chlorobium tepidum.
 OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
 CC Chlorobiaceae.
 CX NCBI_TaxID=1097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TLS / ARCC 49652 / DSM 12025;
 RX MEDLINE=22103685; PubMed=12093901;
 RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
 Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
 Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,

RA Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
 RA Nieman W.C., Feldblum T.V., Hansen C.L., Craven M.B., Radune D.,
 RA Vamathavan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
 RA Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.,
 RT "The complete genome sequence of *Chlorobium tepidum* TLS, a
 RT photosynthetic, anaerobic, green-sulfur bacterium.",
 RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514 (2002).
 DR EMBL; AF012849; AAM72050.1; -.
 DR TIGR; CT0814; -.
 KW Hypochemical protein; Complete proteome.
 SQ SEQUENCE 251 AA; 27932 MW; F7FDA5BD4FCF7985 CRC64;

Query Match 76.1%; Score 35; DB 16; Length 251;
 Best Local Similarity 66.7%; Pred. No. 81;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 LSLIMMIT 9
 DB 85 ELSLIMMIT 93

RESULT 10

O9TC86 PRELIMINARY; PRT; 311 AA.
 AC O9TC86;
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-Oct-2002 (TREMBlrel. 22, Last annotation update)
 DE NADH dehydrogenase subunit 1 (EC 1.6.5.3) (NADH-ubiquinone
 DE oxidoreductase chain 1).
 GN NDI.
 OS Daphnia pulex.
 OS Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Diplostercia;
 OC Cladocera; Anomopoda; Daphniidae; Daphnia.
 OC NCBI_TaxID=6669;
 RN [1]
 RP SEQUENCE FROM N.A.

RA "The complete sequence and gene organization of the mitochondrial
 RT genome of the branchiopod crustacean, *Daphnia pulex*.";
 CC Gene 0:0-0(1999).
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
 CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.
 DR EMBL; AF117817; AAD3342.1; -.
 DR InterPro; IPR001694; Resp_NADH_dhl.
 DR Pfam; PF00146; NADHdh; 1.
 DR PROSITE; PS00667; COMPLEX1_NDI_1; 1.
 DR PROSITE; PS00668; COMPLEX1_NDI_2; 1.
 KW NAD; Oxidoreductase; Transmembrane; Ubiquinone; Mitochondrion.
 SQ SEQUENCE 311 AA; 34610 MW; E1A5B734D954D5C CRC64;

Query Match 76.1%; Score 35; DB 8; Length 311;
 Best Local Similarity 75.0%; Pred. No. 98;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 LSLIMMIT 9
 DB 82 LSLIMMIT 89

RESULT 11

O9I337 PRELIMINARY; PRT; 349 AA.
 AC O9I337;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Translocation protein in type III secretion.
 GN PSTU OR PA1690.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.

OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kae A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lozy S., Olson M.V.,
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964 (2000).
 DR EMBL; AB004596; AAG05079.1; -.
 DR InterPro; IPR006135; Bac_Export_2.
 DR InterPro; IPR006307; HRPV_HrcU.
 DR Pfam; PF01312; Bac_export_2; 1.
 DR PRINTS; PR00950; TYPE3IMSBROT.
 DR TIGRFAWS; TIGR01404; RihB_rel_III; 1.
 KW Complete proteome.
 SQ SEQUENCE 349 AA; 39152 MW; 473C8F0ED5913B61 CRC64;

Query Match 76.1%; Score 35; DB 16; Length 349;
 Best Local Similarity 75.0%; Pred. No. 11e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 LSLIMMIT 9
 DB 150 LSLIMMIT 157

RESULT 12

O9SUFO PRELIMINARY; PRT; 368 AA.
 AC O9SUFO;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-Oct-2002 (TREMBlrel. 22, Last annotation update)
 DE Nodulin-like protein.
 GN T12G13.140 OR A49G08300.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Euphorbia; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Euphorbia II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Bevan M., Lennard N., Quail M., Harris B., Rajandream M.A.,
 RA Barrell B.G., Bancroft I., Newes H.W., Mayer K.F.X., Lemcke K.,
 RA Schueller C.,
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.

RA Lennard N., Quail M., Harris B., Rajandream M.A., Barrell B.G.,
 RA Mewes H.W., Lemcke K., Mayer K.F.X.,
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL080252; CAB45800.1; -.
 DR EMBL; AL161511; CAB7995.1; -.
 DR InterPro; IPR000620; DUF6.
 DR Pfam; PF00892; DUF6; 2.
 SQ SEQUENCE 368 AA; 40015 MW; 62F9270D989C19CF CRC64;

Query Match 76.1%; Score 35; DB 10; Length 368;
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;

Matches 6, Conservative 2, Mismatches 0, Indels 0, Gaps 0;
 QY 1 QLSLMMI 8
 :||:|
 Db 211 ELSLMMI 218

RESULT 13

O41942 PRELIMINARY; PRT; 717 AA.

AC 041942;
 DT 01-JAN-1998 (TRENBLREL. 05, Created)
 DT 01-JAN-1998 (TRENBLREL. 05, Last sequence update)
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
 DE Hypothetical 82.9 kDa protein.
 GN GAMMAY. ORF24.
 OS Murid herpesvirus 4
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae.
 OC NCBI_TaxID=33708;
 OX [1]
 RN SEQUENCE FROM N.A.

RC STRAIN=MTMS and G2.4;
 RX MEDLINE=9736649; PubMed=9223479;
 RA Valgin H.W. IV, Lactelle P., Mamsley P., Hallsworth K., Neck K.E.,
 Dal Canto A.J., Speck S.H.;
 RT "Complete sequence and genomic analysis of murine gammaherpesvirus
 68.";
 RL J. Virol. 71:5894-5904(1997).

[2]
 SEQUENCE FROM N.A.

RC STRAIN=MTMS;
 RA Lactelle P., Mamsley P., Waterston R.H.;
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

[3]
 SEQUENCE FROM N.A.

RA Milligan S., Efethiou S., Stewart J.P., Nash A.A., Davison A.J.;
 RT "Genetic content of murine gammaherpesviruses.";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U97553; AAB66434.1;
 DR EMBL; AF105037; AAF19289.1;
 DR InterPro: IPR004285; U887.
 DR Pfam: PF03043; U887.1.
 KM Hypothetical protein.
 SQ SEQUENCE 717 AA; 82901 MW; 51EF2DC0D59D440 CRC64;

Query Match 76.1%; Score 35; DB 12; Length 717;
 Best Local Similarity 66.7%; Pred. No. 2e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QLSLMMI 9
 :||:|
 Db 25 QLSLMMI 33

RESULT 14

O8L7F7 PRELIMINARY; PRT; 739 AA.

AC 08L7F7;
 DT 01-OCT-2002 (TRENBLREL. 22, Created)
 DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
 DE Hypothetical protein.
 GN AT2G45005.

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Yamada K., Chan M.M., Chang C.H., Dale J.N., Deng J.M., Hsuan V.W.,

Lee J.M., Quach H.L., Tang C., Toriumi M., Wallender E.K., Wong C.,
 Wu H.C., Yu G., Yuan S., Bowser L., Carlini P., Chen H., Cheuk R.,
 Hayashizaki Y., Ishida J., Jones T., Kamita A., Karlin-Neumann G.,
 Kawai J., Kim C., Lam B., Lin J., Miranda M., Narusaka M., Nguyen M.,
 Palm C.J., Sakurai T., Satou M., Seki M., Shin P., Southwick A.,
 Shinzaki K., Davis R.W., Ecker J.R., Theologis A.,
 RT "Arabidopsis Open Reading Frame (ORF) Clones.";
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY133874; AAM91808.1;
 KM Hypothetical protein.
 SQ SEQUENCE 739 AA; 73481 MW; CD4DEAD5C786357 CRC64;

Query Match 76.1%; Score 35; DB 10; Length 739;
 Best Local Similarity 87.5%; Pred. No. 2.1e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QLSLMMI 8
 :||:|
 Db 699 QLSLMMI 706

RESULT 15

O8FKE7 PRELIMINARY; PRT; 64 AA.

AC 08FKE7;
 DT 01-MAR-2003 (TRENBLREL. 23, Created)
 DT 01-MAR-2003 (TRENBLREL. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
 DE Hypothetical protein.
 GN C0481.

OS Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OC NCBI_TaxID=217992;
 OX [1]
 RN SEQUENCE FROM N.A.

RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
 RX MEDLINE=22388234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G., III, Redford P., Roesch P.,
 Raske D., Buckles E.L., Lion S.-R., Boutin A., Hackett J., Stroud D.,
 Mayne G.F., Rose D.O., Zhou S., Schwartz D.C., Perna N.T.,
 RT "Extensive mosaic structure revealed by the complete genome sequence
 of uropathogenic Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 DR EMBL; AE016756; AAN78959.1;
 KM Hypothetical protein; Complete proteome.

SQ SEQUENCE 64 AA; 7237 MW; 53F4B1EF1034F9A3 CRC64;

Query Match 73.9%; Score 34; DB 16; Length 64;
 Best Local Similarity 55.6%; Pred. No. 38;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 QLSLMMI 9
 :||:|
 Db 33 ELSLMMI 41

Search completed: February 10, 2004, 17:19:46
 Job time : 28.069 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 17:19:51 ; Search time 25.1379 Seconds
(without alignments)
74.964 Million cell updates/sec

Title: US-10-023-182-6
Perfect score: 46
Sequence: 1 QLSLMMIT 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09C_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10C_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	9	9	US-09-751-798-6
2	46	100.0	9	12	US-09-766-889A-52
3	46	100.0	9	12	US-10-170-832-61
4	46	100.0	9	12	US-09-833-039-122
5	46	100.0	9	12	US-10-239-313A-466
6	46	100.0	9	12	US-10-364-614-11
7	46	100.0	9	12	US-10-177-277-122
8	46	100.0	9	12	US-10-177-277-132
9	46	100.0	9	12	US-10-023-182-6
10	46	100.0	9	15	US-10-080-013-26
11	46	100.0	9	15	US-10-161-097-25
12	46	100.0	28	12	US-10-117-937-254
13	46	100.0	28	12	US-10-117-937-255
14	46	100.0	135	12	US-10-295-027-388
15	46	100.0	180	9	US-09-751-798-6

16	46	100.0	180	12	US-09-849-602-30	Sequence 30, Appl
17	46	100.0	180	12	US-10-026-066-3	Sequence 3, Appl
18	46	100.0	180	12	US-10-117-937-74	Sequence 74, Appl
19	46	100.0	180	12	US-10-117-937-75	Sequence 75, Appl
20	46	100.0	180	12	US-10-295-027-386	Sequence 386, App
21	46	100.0	180	14	US-10-023-182-8	Sequence 8, Appl
22	46	100.0	180	15	US-10-146-473-69	Sequence 69, Appl
23	46	100.0	180	15	US-10-207-655-71	Sequence 71, Appl
24	46	100.0	397	9	US-09-821-883-27	Sequence 27, Appl
25	41	89.1	9	12	US-10-117-937-199	Sequence 199, App
26	39	84.8	49	12	US-10-144-929-244	Sequence 244, App
27	39	84.8	49	15	US-10-144-928-244	Sequence 244, App
28	39	84.8	154	12	US-10-144-928-240	Sequence 240, App
29	39	84.8	154	15	US-10-144-929-240	Sequence 240, App
30	39	84.8	160	12	US-10-264-237-1525	Sequence 1525, Ap
31	39	84.8	180	12	US-10-364-614-14	Sequence 14, Appl
32	38	82.6	212	12	US-10-002-631C-182	Sequence 182, App
33	37	80.4	9	9	US-09-751-798-5	Sequence 5, Appl
34	37	80.4	9	10	US-09-766-889A-51	Sequence 51, Appl
35	37	80.4	9	12	US-10-170-832-80	Sequence 80, Appl
36	37	80.4	9	12	US-10-364-614-10	Sequence 10, Appl
37	37	80.4	9	12	US-10-177-277-131	Sequence 131, App
38	37	80.4	9	12	US-10-117-937-197	Sequence 197, App
39	37	80.4	9	12	US-10-164-078A-14	Sequence 14, Appl
40	37	80.4	9	12	US-10-164-121A-15	Sequence 15, Appl
41	37	80.4	9	14	US-10-023-182-5	Sequence 5, Appl
42	37	80.4	9	15	US-10-274-017-1	Sequence 1, Appl
43	37	80.4	9	15	US-10-080-013-24	Sequence 24, Appl
44	37	80.4	9	15	US-10-080-013-25	Sequence 25, Appl
45	37	80.4	9	15	US-10-161-097-34	Sequence 34, Appl

ALIGNMENTS

RESULT 1
US-09-751-798-6
Sequence 6, Application US/09751798
Patent No. US20020010321A1
GENERAL INFORMATION:
APPLICANT: Stockert, Elisabeth; Uger, Elke;
APPLICANT: Chen, Yao-tseng; Scanlan, Matthew;
APPLICANT: Knuth, Alexander; Old, Lloyd J.
TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer
TITLE OF INVENTION: Associated Proteins, Uses Thereof,
TITLE OF INVENTION: Truncated Forms of NY-ESO-1, and HLA
TITLE OF INVENTION: Binding Peptides Derived Therefrom
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fulbright & Jaworski, L.L.P.
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 KB storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/751,798
FILING DATE:
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: 09/062,422
FILING DATE:
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: 08/725,182
FILING DATE: October 3, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. US20020010321A1man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5466.3

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3168
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-751-798-6

Query Match 100.0%; Score 46; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QLSLWMT 9
DB 1 QLSLWMT 9

RESULT 2
US-09-766-889A-52
Sequence 52, Application US/09766889A
Patent No. US2002016465A1
GENERAL INFORMATION:
APPLICANT: Luiten, Rosalie
APPLICANT: Boon-Falleur, Thierry
APPLICANT: Van der Bruggen, Pierre
APPLICANT: Stroobant, Vincent
APPLICANT: Demotte, Nathalie
APPLICANT: Schultz, Erwin
TITLE OF INVENTION: MAGE ANTIGENIC PEPTIDES WHICH BIND HLA-B35 AND HLA-B44
FILE REFERENCE: L0461/7104
CURRENT APPLICATION NUMBER: US/09/766,889A
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: US 60/177,242
PRIOR FILING DATE: 2000-01-20
PRIOR APPLICATION NUMBER: US 60/243,212
PRIOR FILING DATE: 2000-10-25
NUMBER OF SEQ ID NOS: 59
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 52
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
US-09-766-889A-52

Query Match 100.0%; Score 46; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLSLWMT 9
DB 1 QLSLWMT 9

RESULT 3
US-10-170-832-81
Sequence 81, Application US/10170832
Publication No. US2003010792A1
GENERAL INFORMATION:
APPLICANT: Chaux, Pascal
APPLICANT: Vantomme, Valérie
APPLICANT: Stroobant, Vincent
APPLICANT: Boon-Falleur, Thierry
APPLICANT: Van der Bruggen, Pierre
APPLICANT: Thielemans, Kris
APPLICANT: Cortals, Jürgen
TITLE OF INVENTION: MAGE-3 PEPTIDES PRESENTED BY HLA CLASS II MOLECULES
FILE REFERENCE: L0461/7052
CURRENT APPLICATION NUMBER: US/10/170,832
PRIOR FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: US/09/166,448
PRIOR FILING DATE: 1998-10-05

NUMBER OF SEQ ID NOS: 81
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 81
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
US-10-170-832-81

Query Match 100.0%; Score 46; DB 12; Length 9;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QLSLWMT 9
DB 1 QLSLWMT 9

RESULT 4
US-09-833-039-122
Sequence 122, Application US/09833039
Publication No. US20030175960A1
GENERAL INFORMATION:
APPLICANT: Tureci, Ozlem
APPLICANT: Sahin, Ugur
APPLICANT: Pfreundschuh, Michael
TITLE OF INVENTION: Tumor Associated Peptide and Uses Thereof
FILE REFERENCE: LUD 5622.1
CURRENT APPLICATION NUMBER: US/09/833,039
PRIOR FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: US 09/409,455
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: US 09/344,040
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 09/105,839
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: US 08/851,130
PRIOR FILING DATE: 1997-05-05
NUMBER OF SEQ ID NOS: 129
SEQ ID NO 122
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
US-09-833-039-122

Query Match 100.0%; Score 46; DB 12; Length 9;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLSLWMT 9
DB 1 QLSLWMT 9

RESULT 5
US-10-239-313A-466
Sequence 466, Application US/10239313A
Publication No. US20030175285A1
GENERAL INFORMATION:
APPLICANT: KLINGNER - HAMOUR, Christine
APPLICANT: CORVAT, Nathalie
APPLICANT: BECK, Alain
APPLICANT: GOETSCH, Liliane
TITLE OF INVENTION: MOLECULE OF PHARMACEUTICAL INTEREST COMPRISING AT ITS N-TERMINAL A GLUTAMIC ACID OR A GLUTAMINE IN THE FORM
FILE REFERENCE: 343 727 - US
CURRENT APPLICATION NUMBER: US/10/239,313A
PRIOR FILING DATE: 2002-09-19
PRIOR APPLICATION NUMBER: FR 00/03711
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT 01/70772
PRIOR FILING DATE: 2001-03-22
NUMBER OF SEQ ID NOS: 697

SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 466
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
US-10-239-313A-466

Query Match
Best Local Similarity 100.0%; Score 46; DB 12; Length 9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLSLMMIT 9
Db 1 QLSLMMIT 9

RESULT 6
US-10-364-614-11
Sequence 11, Application US/10364614
Publication No. US20030175250A1
GENERAL INFORMATION:
APPLICANT: JAGER, Elke
APPLICANT: KNUTH, Alexander
APPLICANT: OLD, Lloyd
APPLICANT: Gnifatic, Sacha
TITLE OF INVENTION: ISOLATED PEPTIDES WHICH BIND TO HLA MOLECULES AND USES THEREOF
FILE REFERENCE: LUD 5726.1 CIP
CURRENT APPLICATION NUMBER: US/10/364,614
CURRENT FILING DATE: 2003-02-24
PRIOR APPLICATION NUMBER: US 607355, 828
PRIOR FILING DATE: 2002-02-13
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn version 3.2
SEQ ID NO 11
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
US-10-364-614-11

Query Match
Best Local Similarity 100.0%; Score 46; DB 12; Length 9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLSLMMIT 9
Db 1 QLSLMMIT 9

RESULT 7
US-10-177-277-122
Sequence 122, Application US/10177277
Publication No. US20030185844A1
GENERAL INFORMATION:
APPLICANT: Tureci, Ozlem
APPLICANT: Sahin, Ugur
APPLICANT: Pfeundschn, Michael
APPLICANT: Ramensee, Hans Georg
APPLICANT: Stevanovic, Stefan
TITLE OF INVENTION: Method for Determining Presence of Cancer in a Sample By Determining Expression of an SSX Gene, Peptides Derived From Said SSX Gene
FILE REFERENCE: LUD 5556.1
CURRENT APPLICATION NUMBER: US/10/177,277
CURRENT FILING DATE: 2002-06-21
PRIOR APPLICATION NUMBER: US/09/344,040
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 09/105,839
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: US 08/851,130
PRIOR FILING DATE: 1997-05-05
NUMBER OF SEQ ID NOS: 132
SEQ ID NO 122
LENGTH: 9

TYPE: PRT
ORGANISM: Homo sapiens
US-10-177-277-122

Query Match
Best Local Similarity 100.0%; Score 46; DB 12; Length 9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLSLMMIT 9
Db 1 QLSLMMIT 9

RESULT 8
US-10-177-277-132
Sequence 132, Application US/10177277
Publication No. US20030185844A1
GENERAL INFORMATION:
APPLICANT: Tureci, Ozlem
APPLICANT: Sahin, Ugur
APPLICANT: Pfeundschn, Michael
APPLICANT: Ramensee, Hans Georg
APPLICANT: Stevanovic, Stefan
TITLE OF INVENTION: Method for Determining Presence of Cancer in a Sample By Determining Expression of an SSX Gene, Peptides Derived From Said SSX Gene
FILE REFERENCE: LUD 5556.1
CURRENT APPLICATION NUMBER: US/10/177,277
CURRENT FILING DATE: 2002-06-21
PRIOR APPLICATION NUMBER: US/09/344,040
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 09/105,839
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: US 08/851,130
PRIOR FILING DATE: 1997-05-05
NUMBER OF SEQ ID NOS: 132
SEQ ID NO 132
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
US-10-177-277-132

Query Match
Best Local Similarity 100.0%; Score 46; DB 12; Length 9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLSLMMIT 9
Db 1 QLSLMMIT 9

RESULT 9
US-10-023-182-6
Sequence 6, Application US/10023182
Publication No. US20020164665A1
GENERAL INFORMATION:
APPLICANT: Stockert, Elisabeth; Jager, Elke;
Chen, Yao-tseng; Scanlan, Matthew;
Knaub, Alexander; Old, Lloyd J.
TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer Associated Proteins, Uses Thereof, Truncated Forms of NY-ESO-1, and HLA Binding Peptides Derived Therefrom
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESS: Fulbright & Jaworski, L.L.P.
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 Kb storage


```
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/10/023,182
  FILING DATE: 17-Dec-2001
  PRIORITY APPLICATION DATA:
    APPLICATION NUMBER: 09/751,798
    FILING DATE: December 29, 2000
    APPLICATION NUMBER: 09/062,422
    FILING DATE: <unknown>
    APPLICATION NUMBER: 08/725,182
    FILING DATE: October 3, 1996
  ATTORNEY/AGENT INFORMATION:
    NAME: Hanson, No. US20020164665A1man D.
    REGISTRATION NUMBER: 30,946
    REFERENCE/DOCKET NUMBER: LUD 5466.3
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (212) 318-3168
    TELEFAX: (212) 752-5958
  INFORMATION FOR SEQ ID NO: 6:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 9 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
  SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-023-182-6
```

```
Query Match      100.0%; Score 46; DB 14; Length 9;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 QLSLWMIT 9
Db 1 QLSLWMIT 9
```

```
RESULT 10
US-10-080-013-26
; Sequence 26, Application US/10080013
; Publication No. US20030077246A1
; GENERAL INFORMATION:
; APPLICANT: Moriarty, Ann
; APPLICANT: Leturcq, Didier
; APPLICANT: Degraw, Juli
; APPLICANT: Helskala, Marja
; APPLICANT: Peterson, Per
; APPLICANT: Jackson, Michael
; TITLE OF INVENTION: A CELL THERAPY METHOD FOR THE TREATMENT OF TUMORS
; FILE REFERENCE: OMT-1557
; CURRENT APPLICATION NUMBER: US/10/080,013
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-080-013-26
```

```
Query Match      100.0%; Score 46; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 QLSLWMIT 9
Db 1 QLSLWMIT 9
```

```
RESULT 11
US-10-161-097-35
; Sequence 35, Application US/10161097
; Publication No. US20030096404A1
```

```
GENERAL INFORMATION:
; APPLICANT: ROSENZWEIG, Michael
; APPLICANT: PYKEET, Mark J.
; APPLICANT: SCADDEN, David T.
; APPLICANT: KOZMANSKY, Mark C.
; TITLE OF INVENTION: LYMPHOID TISSUE-SPECIFIC CELL PRODUCTION
; TITLE OF INVENTION: FROM HEMATOPOIETIC PROGENITOR CELLS IN THREE-DIMENSIONAL
; FILE REFERENCE: C1005/7012/KX/ERG
; CURRENT APPLICATION NUMBER: US/10/161,097
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US/09/574,749
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US 60/107,972
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: PCT/US99/26795
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: US 09/524,749
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 35
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Homo Sapiens source
US-10-161-097-35
```

```
Query Match      100.0%; Score 46; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 QLSLWMIT 9
Db 1 QLSLWMIT 9
```

```
RESULT 12
US-10-117-937-254
; Sequence 254, Application US/10117937
; Publication No. US2003020239A1
; GENERAL INFORMATION:
; APPLICANT: CTL IMMUNO THERAPIES CORP.
; APPLICANT: SIMARD, John, J.L.
; APPLICANT: DIAMOND, David, C.
; APPLICANT: Liu, Liping
; APPLICANT: XIE, Zhidong
; TITLE OF INVENTION: EPITOPE SEQUENCES
; FILE REFERENCE: CTLIMM.027A
; CURRENT APPLICATION NUMBER: US/10/117,937
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/282,211
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/337,017
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/363,210
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 602
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 254
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-937-254
```

```
Query Match      100.0%; Score 46; DB 12; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 QLSLWMIT 9
Db 20 QLSLWMIT 28
```

RESULT 13
US-10-117-937-255
; Sequence 255, Application US/10117937
; Publication No. US20030220239A1
; GENERAL INFORMATION:
; APPLICANT: CTL IMMUNO THERAPIES CORP.
; APPLICANT: SIMARD, John, J.L.
; APPLICANT: DIAMOND, David, C.
; APPLICANT: LIU, Liping
; APPLICANT: XIE, Zhidong
; TITLE OF INVENTION: EPITOPE SEQUENCES
; FILE REFERENCE: CTILIM.027A
; CURRENT APPLICATION NUMBER: US/10/117,937
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/282,211
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/337,017
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/363,210
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 602
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 255
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-937-255

Query Match
Best Local Similarity 100.0%; Score 46; DB 12; Length 28;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLSLWMIT 9
Db 6 QLSLWMIT 14

RESULT 14
US-10-295-027-388
; Sequence 388, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afari, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08

PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 388
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-388

Query Match
Best Local Similarity 100.0%; Score 46; DB 12; Length 135;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLSLWMIT 9
Db 110 QLSLWMIT 118

RESULT 15
US-09-751-798-8
; Sequence 8, Application US/09751798
; Patent No. US20020010321A1
; GENERAL INFORMATION:
; APPLICANT: Stockert, Elisabeth; Jager, Elke;
; APPLICANT: Chen, Yao-tseng; Scanlan, Matthew;
; APPLICANT: Knuth, Alexander; Old, Lloyd J.
; TITLE OF INVENTION: Antibodies which Bind to NY-ESO-1 Cancer
; TITLE OF INVENTION: Truncated Forms of NY-ESO-1, and HLA
; TITLE OF INVENTION: Binding Peptides Derived Therefrom
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pulbright & Jaworski, L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/751,798
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/062,422
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 06/725,182
; FILING DATE: October 3, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. US20020010321A1man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5466.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3168
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180
; TYPE: amino acid
; TOPOLOGY: linear
US-09-751-798-8

Query Match
100.0%; Score 46; DB 9; Length 180;

Fri Feb 13 14:37:41 2004

us-10-023-182-6.rapb

Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLSLMMIT 9
|||
Db 155 QLSLMMIT 163

Search completed: February 10, 2004, 17:32:13
Job time : 25.1379 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 17:18:11 ; Search time 13.348 Seconds
(without alignments)
53.077 Million cell updates/sec

Title: US-10-023-182-6
Perfect score: 46
Sequence: 1 QUSLMMWT 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 378584 seqs, 7869973 residues

Total number of hits satisfying chosen parameters: 378584

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_AA.New:*
1: /cgn2_6/prodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/prodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	9	5	US-09-165-546D-6
2	46	100.0	9	6	US-10-706-475-3
3	46	100.0	9	6	US-10-289-566-26
4	46	100.0	9	6	US-10-751-088-6
5	46	100.0	30	6	US-10-296-734-1416
6	46	100.0	30	6	US-10-296-734-1418
7	46	100.0	30	6	US-10-296-734-1448
8	46	100.0	30	6	US-10-296-734-1450
9	46	100.0	135	1	PCT-US02-29560A-330
10	46	100.0	180	1	PCT-US02-29560A-129
11	46	100.0	180	1	PCT-US03-27706-74
12	46	100.0	180	1	PCT-US03-27706-75
13	46	100.0	180	1	PCT-US03-28193-1270
14	46	100.0	180	1	PCT-US04-00885-95
15	46	100.0	180	5	US-09-165-546D-15
16	46	100.0	180	6	US-10-296-734-832
17	46	100.0	180	6	US-10-296-734-834
18	46	100.0	180	6	US-10-657-022-74
19	46	100.0	180	6	US-10-657-022-75
20	46	100.0	180	6	US-10-723-860-1270
21	46	100.0	180	6	US-10-751-088-15
22	46	100.0	180	6	US-10-756-149-5024
23	46	100.0	3541	6	US-10-296-734-1454
24	37	80.4	9	1	PCT-US03-27125-25
25	37	80.4	9	5	US-09-165-546D-5
26	37	80.4	9	6	US-10-651-616-25

27	37	80.4	9	6	US-10-706-475-2	Sequence 2, Appl
28	37	80.4	9	6	US-10-706-475-6	Sequence 6, Appl
29	37	80.4	9	6	US-10-706-475-7	Sequence 7, Appl
30	37	80.4	9	6	US-10-706-475-8	Sequence 8, Appl
31	37	80.4	9	6	US-10-706-475-10	Sequence 10, Appl
32	37	80.4	9	6	US-10-706-475-12	Sequence 12, Appl
33	37	80.4	9	6	US-10-289-566-24	Sequence 24, Appl
34	37	80.4	9	6	US-10-289-566-25	Sequence 25, Appl
35	37	80.4	9	6	US-10-751-088-5	Sequence 5, Appl
36	37	80.4	10	6	US-10-706-475-11	Sequence 11, Appl
37	37	80.4	10	6	US-10-706-475-13	Sequence 13, Appl
38	37	80.4	10	6	US-10-706-475-14	Sequence 14, Appl
39	37	80.4	10	6	US-10-706-475-15	Sequence 15, Appl
40	37	80.4	10	6	US-10-706-475-16	Sequence 16, Appl
41	37	80.4	10	6	US-10-706-475-17	Sequence 17, Appl
42	37	80.4	10	6	US-10-706-475-18	Sequence 18, Appl
43	37	80.4	11	5	US-09-165-546D-4	Sequence 4, Appl
44	37	80.4	11	6	US-10-706-475-1	Sequence 1, Appl
45	37	80.4	11	6	US-10-751-088-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-165-546D-6
Sequence 6, Application US/09165546D
GENERAL INFORMATION:
APPLICANT: Knuth, Alexander; Jager, Elke; Chen, Scanlan, Matt;
Gure, Ali, Old, Lloyd, Ritzer, Gerd
TITLE OF INVENTION: ISOLATED PEPTIDES CORRESPONDING TO AMINO ACID
SEQUENCES OF NY-ESO-1, WHICH BIND TO MHC CLASS I AND MHC
US05 THEREOF
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: FUTURIGHT & JAWORSKI LLP
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10158
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/165,546D
FILING DATE: 02-Oct-1998
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/062,422
FILING DATE: April 17, 1998
APPLICATION NUMBER: 08/937,263
FILING DATE: September 15, 1997
APPLICATION NUMBER: US 08/725,182
FILING DATE: October 3, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, Norman D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 2166.4 CIP (09807811)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3000
TELEFAX: (212) 318-3400
INFORMATION FOR SEQ. ID NO. 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-165-546D-6
Query Match 100.0%; Score 46; DB 5; Length 9;

Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLSLMMIT 9
Db 1 QLSLMMIT 9

RESULT 2
US-10-706-475-3

; Sequence 3, Application US/10706475
; GENERAL INFORMATION:
; APPLICANT: Cerundolo, Vincenzo
; TITLE OF INVENTION: Derivatives of NY-ESO-1 Peptides and Uses Thereof
; FILE REFERENCE: LDD 5625.1
; CURRENT FILING DATE: 2003-11-11
; PRIOR APPLICATION NUMBER: US/09/514,036E
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: US 09/440,621
; PRIOR FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: US 09/165,546
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: US 09/062,422
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: US 08/725,162
; PRIOR FILING DATE: 1996-10-03
; NUMBER OF SEQ ID NOS: 18
; SEQ ID NO 3
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-706-475-3

Query Match 100.0%; Score 46; DB 6; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLSLMMIT 9
Db 1 QLSLMMIT 9

RESULT 3
US-10-289-566-26
; Sequence 26, Application US/10289566
; GENERAL INFORMATION:
; APPLICANT: Leturcq, Didier J.
; APPLICANT: Moriarty, Ann M.
; APPLICANT: Jackson, Michael R.
; APPLICANT: Peterson, Per A.
; APPLICANT: Richards, Jon M.
; TITLE OF INVENTION: A CELL THERAPY METHOD FOR THE TREATMENT OF TUMORS
; FILE REFERENCE: ORT 1342CIP
; CURRENT FILING DATE: US/10/289,566
; PRIOR APPLICATION NUMBER: 2002-11-07
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: US 60/270,252
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US 10/080,013
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-289-566-26

Query Match 100.0%; Score 46; DB 6; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLSLMMIT 9
Db 1 QLSLMMIT 9

RESULT 4
US-10-751-088-6

; Sequence 6, Application US/10751088
; GENERAL INFORMATION:
; APPLICANT: Knuth, Alexander; Jager, Elke; Chen, Yao; Scanlan, Matt;
; TITLE OF INVENTION: ISOLATED PEPTIDES CORRESPONDING TO AMINO ACID
; SEQUENCES OF NY-ESO-1, WHICH BIND TO MHC CLASS I AND MHC CL
; USES THEREOF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FULBRIGHT & JAWORSKI LLP
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10158
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/751,088
; FILING DATE: 02-Jan-2004
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/165,546D
; FILING DATE: 02-Oct-1998
; APPLICATION NUMBER: 09/062,422
; FILING DATE: April 17, 1998
; APPLICATION NUMBER: 08/937,263
; FILING DATE: September 15, 1997
; APPLICATION NUMBER: US 08/725,182
; FILING DATE: October 3, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, Norman D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LDD 2166.4 CIP (09807811)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 318-3400
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-751-088-6

Query Match 100.0%; Score 46; DB 6; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLSLMMIT 9
Db 1 QLSLMMIT 9

RESULT 5
US-10-296-734-1416
; Sequence 1416, Application US/10296734
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine

```
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU PQ761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1416
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: NYSOLA segment 10
; US-10-296-734-1416

Query Match
Best Local Similarity 100.0%; Score 46; DB 6; Length 30;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLSLMMIT 9
Db 22 QLSLMMIT 30

RESULT 6
US-10-296-734-1418
; Sequence 1418, Application US/10296734
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU PQ761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1418
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: NYSOLA segment 11
; US-10-296-734-1418

Query Match
Best Local Similarity 100.0%; Score 46; DB 6; Length 30;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLSLMMIT 9
Db 7 QLSLMMIT 15

RESULT 7
US-10-296-734-1448
; Sequence 1448, Application US/10296734
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU PQ761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1448
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
```

```
; OTHER INFORMATION: LAGE1 segment 10
; US-10-296-734-1448

Query Match
Best Local Similarity 100.0%; Score 46; DB 6; Length 30;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLSLMMIT 9
Db 22 QLSLMMIT 30

RESULT 8
US-10-296-734-1450
; Sequence 1450, Application US/10296734
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU PQ761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1450
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: LAGE1 segment 11
; US-10-296-734-1450

Query Match
Best Local Similarity 100.0%; Score 46; DB 6; Length 30;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLSLMMIT 9
Db 7 QLSLMMIT 15

RESULT 9
PCT-US02-29560A-330
; Sequence 330, Application PC/TUS0229560A
; GENERAL INFORMATION:
; APPLICANT: Afari, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Gish, Kurt C.
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Wilson, Keith E.
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-002710PC
; CURRENT APPLICATION NUMBER: PCT/US02/29560A
; CURRENT FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: US 60/323,469
; PRIOR FILING DATE: 2001-09-17
; NUMBER OF SEQ ID NOS: 412
; SOFTWARE: FastSeq For Windows Version 3.0
; SEQ ID NO 330
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
; PCT-US02-29560A-330

Query Match
Best Local Similarity 100.0%; Score 46; DB 1; Length 135;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 QSLMWT 9
| | | | |
Db 110 QSLMWT 118

RESULT 10
PCT-US02-29560A-329
; Sequence 329, Application PC/TUS0229560A
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Gish, Kurt C.
; APPLICANT: Hevazi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Wilson, Keith E.
; APPLICANT: Zlotnik, Albert
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-002710PC
; CURRENT APPLICATION NUMBER: PCT/US02/29560A
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: US 60/323,469
; PRIOR FILING DATE: 2001-09-17
; NUMBER OF SEQ ID NOS: 412
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 329
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-29560A-329

Query Match 100.0%; Score 46; DB 1; Length 180;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QSLMWT 9
| | | | |
Db 155 QSLMWT 163

RESULT 11
PCT-US03-27706-74
; Sequence 74, Application PC/TUS0327706
; GENERAL INFORMATION:
; APPLICANT: MANNKIND CORPORATION
; APPLICANT: SIMARD, John J. L.
; APPLICANT: DIAMOND, David C.
; APPLICANT: LIU, Liping
; APPLICANT: LIU, Zheng
; TITLE OF INVENTION: EPIPOPE SEQUENCES
; FILE REFERENCE: MANNK.032VPC
; CURRENT APPLICATION NUMBER: PCT/US03/27706
; PRIOR FILING DATE: 2003-09-05
; PRIOR APPLICATION NUMBER: US 60/409123
; PRIOR FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 610
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-27706-74

Query Match 100.0%; Score 46; DB 1; Length 180;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QSLMWT 9
| | | | |
Db 155 QSLMWT 163

RESULT 12
PCT-US03-27706-75
; Sequence 75, Application PC/TUS0327706
; GENERAL INFORMATION:
; APPLICANT: MANNKIND CORPORATION
; APPLICANT: SIMARD, John J. L.
; APPLICANT: DIAMOND, David C.
; APPLICANT: LIU, Liping
; APPLICANT: LIU, Zheng
; TITLE OF INVENTION: EPIPOPE SEQUENCES
; FILE REFERENCE: MANNK.032VPC
; CURRENT APPLICATION NUMBER: PCT/US03/27706
; PRIOR FILING DATE: 2003-09-05
; PRIOR APPLICATION NUMBER: US 60/409123
; PRIOR FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 610
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-27706-75

Query Match 100.0%; Score 46; DB 1; Length 180;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QSLMWT 9
| | | | |
Db 155 QSLMWT 163

RESULT 13
PCT-US03-38193-1270
; Sequence 1270, Application PC/TUS0338193
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.00PC00
; CURRENT APPLICATION NUMBER: PCT/US03/38193
; PRIOR FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1270
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-38193-1270

Query Match 100.0%; Score 46; DB 1; Length 180;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QSLMWT 9
| | | | |
Db 155 QSLMWT 163

RESULT 14
PCT-US04-00885-95
; Sequence 95, Application PC/TUS0400885
; GENERAL INFORMATION:
; APPLICANT: PDL
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER

```

; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: PCT/US04/00865
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 95
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo Sapiens
; PCT-US04-00865-95

```

```

Query Match      100.0%; Score 46; DB 1; Length 180;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 QUSLWMIT 9
Db      155 QUSLWMIT 163

```

```

RESULT 15
US-09-165-546D-15
; Sequence 15, Application US/09165546D
; GENERAL INFORMATION:
; APPLICANT: Knuth, Alexander; Jager, Elke; Chen, Yao; Scanlan, Matt;
; Gure, Ali, Old, Lloyd, Ritter, Gerd
; TITLE OF INVENTION: ISOLATED PEPTIDES CORRESPONDING TO AMINO ACID
; SEQUENCES OF NY-ESO-1, WHICH BIND TO MHC CLASS I AND MHC CL
; USES THEREOF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FULLRIGHT & JANORSKI LLP
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10158
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/165,546D
; FILING DATE: 02-Oct-1998
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/062,422
; FILING DATE: April 17, 1998
; APPLICATION NUMBER: 08/937,263
; FILING DATE: September 15, 1997
; APPLICATION NUMBER: US 08/725,182
; FILING DATE: October 3, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, Norman D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 2166.4 CIP (09807811)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3800
; TELEFAX: (212) 318-3400
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 15
US-09-165-546D-15

```

```

Query Match      100.0%; Score 46; DB 5; Length 180;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 QUSLWMIT 9

```

```

Db      155 QUSLWMIT 163

```

Search completed: February 10, 2004, 17:30:39
Job time : 13.3448 secs